



Jurnal Informatika

Volume 21, No.1, Juni 2021

Analisis Sentimen Terhadap Kinerja Menteri Kesehatan Indonesia Selama Pandemi Covid-19

Tri Rivanie, Rangga Pebrianto, Taopik Hidayat, Achmad Bayhaqy, Windu Gata, Hafifah Bella Novitasari

Analisa Dan Prediksi Iklan Lowongan Kerja Palsu Dengan Metode *Natural Language Programing* dan *Machine Learning*

Hary Sabita, Fitria, Riko Herwanto

Prediksi Tingkat Kesuksesan Promosi Bank Dengan Algoritma DNN

Oscar, Nurlaelatul Maulidah, Annida Purnamawati, Destiana Putri, Hilman F. Pardede

Prediksi Tingkat Pelanggan *CHURN* Pada Perusahaan Telekomunikasi Dengan Algoritma ADABOOST

Iqbal Muhammad Latief, Agus Subekti, Windu Gata

Penerapan *Data Mining* Terhadap Data Covid-19 Menggunakan Algoritma Klasifikasi

Rizka Dahlia, Nanik Wuryani, Sri Hadiani, Windu Gata, Arina Selawati

Video Rancangan By-Pro (Sistem Pengukuran Dan Pendataan Balita Di Posyandu)

Chyntia Raras Ajeng Widiawati, Indra Alan Nugroho, Irfan Rifai Aziz, Rizky Arsyansyah

Sistem Pakar *Forward Chaining* , *Fuzzy-Max* dan *Certainty Factor* Ayam Pedaging

Asep Afandi, Dwi Marisa Efendi

Pengelolaan Data Produksi Berbasis Web Pada CV. Sumber Proteina Bandar Lampung

TM Zaini, Driya Wiryawan, Aryadi

Augmented Reality Pengenalan Objek Wisata Taman Mumbul Sangheh Berbasis Android

Kadek Agus Wirawan, I Gede Harsemadi, Ni Kadek Sukerti

Audit *E-Learning* Dengan *Framework Cobit 5.0* Di Masa Pandemi Covid-19

Rini Nurlistiani, Neni Purwati, Supri Yanto

Volume 21 Nomor 1, Juni 2021

Jurnal Informatika adalah sarana informasi Ilmu Pengetahuan & Teknologi yang berupa hasil penelitian, studi pustaka maupun tulisan ilmiah. Terbit dalam satu tahun dengan frekuensi pada bulan Juni dan Desember.

Penanggung Jawab : Ir.Firmansyah YA,MBA.,M.Sc

Pembina : Dr. R.Z. Abdul Aziz, ST, M.T

Ketua Penyunting : Dr. Sri Lestari,M.Cs

Sekretaris Penyunting : Suci Mutiara,S.Kom.,M.TI

Penyunting Ahli : Dr. Muhammad Said Hasibuan, M.Kom

Dr. Suhendro Yusuf Irianto, M.Kom

Dr. Sri Lestari, S.Kom., M.Cs

Joko Triloka, PhD

Dr.Wasilah ,MT

Prof. Dr. Arif Muntasa, S.Si., M.T (UniversitasTrunojoyo)

Dr. Slamet Risnanto (Universitas Sangga Buana)

Dr. I Putu Agus Eka Pratama (Universitas Udayana)

Dr. Rangga Firdaus,M.Kom (Universitas Lampung)

Dr. Herri Setiawan (Universitas Indo Global Mandiri)

Dr. Anggoro Suryo Pramudyo (Universitas Sultan Ageng Tirtayasa)

Dr. Triwiyanto (Poltekkes Kemenkes Surabaya,Department of
Electromedical Engineering

Dr. Muhaimin Hasanudin (Universitas Mercuru Buana)

Dr. Ricardus Anggi Pramunendar (Universitas Dian Nuswantoro)

Dr. Rendra Gustriansyah (Universitas Indo Global Mandiri)

Dr.Dewi Agushinta Rahayu,M.Sc (Universitas Gunadarma)

Penyunting Pelaksana

Koordinator : Fitria,ST.,M.Kom

Anggota :

1. M Miftakul Amin,S.Kom.,M.Eng (Poltek Unsri)
2. Yulmaini,S.Kom.,M.CS (IIB Darmajaya)
3. Rio Kurniawan, M.CS(IIB Darmajaya)

Alamat Redaksi/penerbit :

Informatics & Business Institute Darmajaya

Jl. Zainal Abidin P.A. No. 93 Bandar Lampung 35142. Telp. (0721) 787214

Fax (0721) 700261.

Email : lp4mjurin@gmail.com

Website : <http://jurnal.darmajaya.ac.id/index.php/JurnalInformatika>

DARI REDAKSI

Jurnal Informatika IIB Darmajaya Volume 21 No. 1 juni 2021 ini menyajikan sepuluh judul artikel yang beragam dengan kajian yang berbeda. Pada volume ini, berisi beberapa topik dalam bidang informatika.

Topik-topik tersebut antara lain : Metode Natural Language Programing Dan Machine Learning, Analisis Sentimen, Algoritma DNN, Algoritma ADABOOST, Penerapan *Data Mining*, Rancangan By-Pro, Sistem Pakar Forward Chaining , Fuzzy-Max Dan Certainty, Pengelolaan Data Produksi Berbasis Web, Augmented Reality, Audit E-Learning Dengan Framework Cobit 5.0

Demikian ringkasan beberapa sintesis makalah yang ada pada Volume ini, masih ada lagi makalah-makalah yang belum kami sajikan untuk lebih lengkapnya para pembaca dapat mengeksplor makalah-makalah yang lain. Harapan kami mudah-mudahan semua makalah yang disajikan dapat menambah pengetahuan dan pengamalan para pembaca.

Terakhir kami ucapkan banyak terimakasih kepada para penulis atas kontribusinya dalam volume ini.

Selamat membaca.

Bandar Lampung, Juni 2021

Dewan Redaksi

DAFTAR ISI

Judul Artikel	Halaman
Analisis Sentimen Terhadap Kinerja Menteri Kesehatan Indonesia Selama Pandemi Covid-19 <i>Tri Rivanie, Rangga Pebrianto, Taopik Hidayat, Achmad Bayhaqy, Windu Gata, Hafifah Bella Novitasari</i>	1 - 13
Analisa Dan Prediksi Iklan Lowongan Kerja Palsu Dengan Metode Natural Language Programing Dan Machine Learning <i>Hary Sabita, Fitria, Riko Herwanto</i>	14 - 22
Prediksi Tingkat Kesuksesan Promosi Bank Dengan Algoritma DNN <i>Oscar, Nurlaelatul Maulidah, Annida Purnamawati, Destiana Putri, Hilman F. Pardede</i>	23 - 33
Prediksi Tingkat Pelanggan <i>CHURN</i> Pada Perusahaan Telekomunikasi Dengan Algoritma ADABOOST <i>Iqbal Muhammad Latief, Agus Subekti, Windu Gata</i>	34 - 43
Penerapan <i>Data Mining</i> Terhadap Data Covid-19 Menggunakan Algoritma Klasifikasi <i>Rizka Dahlia, Nanik Wuryani, Sri Hadiani, Windu Gata, Arina Selawati</i>	44 - 52
Video Rancangan By-Pro (Sistem Pengukuran Dan Pendataan Balita Di Posyandu) <i>Chyntia Raras Ajeng Widiawati, Indra Alan Nugroho, Irfan Rifai Aziz, Rizky Arsyansyah</i>	53 - 64
Sistem Pakar Forward Chaining , Fuzzy-Max Dan Certainty Factor Ayam Pedaging <i>Asep Afandi, Dwi Marisa Efendi</i>	65 - 73
Pengelolaan Data Produksi Berbasis Web Pada CV. Sumber Proteina Bandar Lampung <i>TM Zaini, Driya Wiryawan, Aryadi</i>	74 - 83
Augmented Reality Pengenalan Objek Wisata Taman Mumbul Sangeh Berbasis Android <i>Kadek Agus Wirawan, I Gede Harsemadi, Ni Kadek Sukerti</i>	84 - 89
Audit E-Learning Dengan Framework Cobit 5.0 Di Masa Pandemi Covid-19 <i>Rini Nurlistiani, Neni Purwati, Supri Yanto</i>	90 - 103

ANALISIS SENTIMEN TERHADAP KINERJA MENTERI KESEHATAN INDONESIA SELAMA PANDEMI COVID-19

Tri Rivanie¹, Rangga Pebrianto², Taopik Hidayat³, Achmad Bayhaqy⁴, Windu Gata⁵, Hafifah Bella Novitasari⁶

^{1,2,3,4,5,6} Ilmu Komputer, STMIK Nusa Mandiri Jakarta

Jl. Kramat Raya No.18, Jakarta Pusat, Jakarta 10420

Email: 14002326@nusamandiri.ac.id¹, 14002396@nusamandiri.ac.id²,

Taopik.Tao@nusamandiri.ac.id³, achmad.acq@nusamandiri.ac.id⁴, Windu@nusamandiri.ac.id⁵, hafifahbela@stmik-banisaleh.ac.id⁶

ABSTRACT

The pandemic that occurred in Indonesia has not yet subsided and far from under control. Indonesian Ministry of Health is most appropriate person to responsible for providing an explanation of actual situation and extent to which state has handled it. However, he has rarely appeared in public lately to explain about handling of Covid-19 pandemic. In response, many people are pros and cons come to give their opinions and feedback. The increasing use of internet during pandemic, especially on social media, where one of them is Twitter, which is a means of expressing opinions. Posting tweets is a community habit to assess or respond to events, as well as represent public's response to an event, especially Ministry of Health steps and policies in handling and breaking chain of Covid-19 pandemic.

The tweet posts were taken only in Indonesian-language and also related to performance of Government, especially Ministry of Health. After that, a label is given so that sentiment of tweets is known. To test results of these sentiments, an algorithm is used by comparing two methods of Support Vector Machine (SVM) and Naïve Bayes (NB). Validation was carried out using k-Fold Cross Validation to obtain an accuracy value. The results show that accuracy value for NB algorithm is 66.45% and SVM algorithm has a greater accuracy value of 72.57%. So it can be seen that SVM algorithm managed to get the best accuracy value in classifying positive comments and negative comments related to sentiment analysis towards Ministry of Health.

Keywords—Support Vector Machine, Naïve Bayes, Analisis sentimen, K-Fold Cross Validation

ABSTRAK

Pandemi yang terjadi di Indonesia belum juga mereda dan terkendali. Menteri kesehatan Indonesia merupakan sosok yang dianggap paling tepat untuk bertanggung jawab memberi penjelasan mengenai situasi sebenarnya dan sejauh apa penanganan yang sudah dilakukan oleh negara. Namun, beliau sangat jarang tampil akhir-akhir ini di hadapan publik untuk menjelaskan mengenai penanganan pandemi Covid-19. Menanggapi hal tersebut, banyak masyarakat yang pro maupun kontra ikut memberikan pendapat serta tanggapan. Meningkatnya penggunaan internet selama masa pandemi terutama pada media sosial, dimana salah satu wadahnya yaitu *twitter* yang menjadi sarana menuangkan pendapat. Memposting *tweet* merupakan kebiasaan masyarakat untuk menilai atau menanggapi peristiwa, serta merepresentasikan tanggapan masyarakat terhadap suatu peristiwa terutama langkah dan kebijakan Kementerian Kesehatan dalam penanganan dan pemutusan rantai pandemi Covid-19.

Postingan *tweet* tersebut diambil yang berbahasa Indonesia dan juga berkaitan dengan kinerja Pemerintah, khususnya Kementerian Kesehatan. Setelah itu, dilakukan pemberian label agar diketahui sentimen dari *tweet* tersebut. Untuk menguji hasil sentimen tersebut, dilakukan menggunakan algoritma dengan membandingkan dua metode *Support Vector Machine* (SVM) dan *Naïve Bayes* (NB). Dilakukan validasi menggunakan *k-Fold Cross Validation* untuk memperoleh nilai akurasi (*accuracy*). Dari hasilnya menunjukkan bahwa nilai akurasi untuk algoritma NB sebesar 66,45% dan algoritma SVM memiliki nilai akurasi lebih besar yaitu 72,57%. Maka dapat diketahui bahwa algoritma SVM berhasil mendapatkan nilai akurasi yang terbaik dalam mengklasifikasikan komentar positif dan komentar negatif terkait dengan analisis sentimen terhadap Kementerian Kesehatan.

Kata Kunci—*Support Vector Machine, Naïve Bayes, Analisis sentimen, K-Fold Cross Validation*

I. PENDAHULUAN

Saat ini seluruh dunia sedang direpotkan dengan virus baru *Covid-19*, *South China Morning Post* melaporkan kasus *Covid-19* pertama di dunia ditemukan pada 17 November 2019 di Cina. Infeksi ini ditemukan pada pria berusia 55 tahun dari Provinsi *Hubei*. Sementara kasus *Covid-19* pertama di Indonesia diumumkan pada 2 Maret 2020 atau sekitar 4 bulan setelah ditemukan kasus *Covid-19* pertama di Cina. Total kasus corona di Indonesia pada 2 Maret sebanyak dua kasus. Kondisi pandemi di Indonesia saat ini belum juga terkendali, data dan angka jelas menunjukkan itu. Berdasarkan data dari pemerintah di *Covid19.go.id*, kurva kasus positif *Covid-19* di Indonesia terus bergerak naik dan belum ada tanda-tanda penurunan. DKI Jakarta menjadi wilayah dengan kasus positif *Covid-19* terbanyak yakni 50.671,

diikuti Jawa Timur 37.093, Jawa Tengah 16.508 dan Jawa Barat 13.668 kasus.

Dalam hal ini sosok Menteri Kesehatan Dr. dr. Terawan Agus Purhanto, Sp.Rad kembali menjadi sorotan. Menteri Kesehatan dianggap menjadi sosok yang bertanggung jawab dalam memberi penjelasan mengenai situasi sebenarnya dan sejauh apa penanganan yang sudah dilakukan oleh Negara. Karena beliau yang memiliki wewenang akses anggaran dan pemberi arahan. Namun kini Menteri Kesehatan Terawan Agus Purhanto sangat jarang tampil di hadapan publik untuk menjelaskan mengenai penanganan pandemi *Covid-19*. Jarang terlihatnya Menkes Terawan di media mulai jadi sorotan sejak jurnalis sekaligus presenter Najwa Shihab dalam acara "Mata Najwa" edisi "Menanti Terawan" bermonolog dengan kursi kosong seolah-olah ada Terawan yang tengah duduk di sana. Dalam monolognya, Najwa mengungkap kegelisahan masyarakat yang

sudah jarang melihat sosok Menkes Terawan di media sejak pandemi *Covid-19* melanda Indonesia. *Rapid test* yang belum mencapai target dan perlindungan terhadap tenaga kesehatan sebagai garda terdepan belum maksimal, angka kematian tenaga kesehatan sangat tinggi dan masih terus naik, dimana Kementerian kesehatan seharusnya menjadi pelindung dan pembela tenaga kesehatan. Selain itu masih saja ada disparitas antara data pusat dan daerah, padahal data tersebut sangat krusial dalam menentukan kebijakan. Bahkan data menunjukkan bahwa Gedung kementerian menjadi *cluster* perkantoran terbesar di Jakarta. membuat opini masyarakat terhadap Menteri kesehatan bahwa kini beliau telah menyadari bahwa Indonesia telah kecolongan dalam penanganan langkah diawal yang seharusnya bisa lebih tanggap. Menurut Najwa, kehadiran Terawan sangat diperlukan untuk menjelaskan kepada publik terkait situasi pandemi *Covid-19* di Indonesia.

Menanggapi hal tersebut banyak masyarakat yang pro dan kontra ikut memberikan pendapat dan tanggapan menyoroti Kinerja Pemerintah terkait *Covid-19*, terutama kinerja Kementerian kesehatan, kemudian memposting di media sosial. Media sosial telah menjadi bagian dari masyarakat global, di mana

orang-orang dari seluruh dunia membagikan pendapat dan pandangan mereka secara bebas di blog pribadi, *Facebook*, *YouTube*, *Twitter* dan lain-lain. salah satu media sosial yang berkembang pesat penggunaannya di Indonesia adalah *Twitter*. *Twitter* adalah *platform* media sosial dianggap paling populer untuk berbagi masalah politik dan masalah social [1]. Saat ini *Twitter* merupakan sebuah indikator yang baik untuk memberikan pengaruh dalam penelitian [2]. Banyaknya *tweet* dari pengguna *twitter* hari ke hari bertambah banyak, data yang berkembang di *Twitter* dianggap lebih diminati karena dirasa cukup mudah dan simpel dalam merepresentasikan apa yang berkembang di masyarakat pada umumnya.

Informasi yang tersimpan cenderung berupa teks dilakukan proses ekstraksi informasi yang berpola (*text mining*). *Text mining* adalah proses ekstraksi pola (informasi dan pengetahuan yang berguna) dari sejumlah besar sumber data tak terstruktur. *Text mining* memiliki tujuan dan menggunakan proses yang sama dengan *Data Mining*, namun memiliki masukan yang berbeda. Masukan untuk *text mining* adalah data yang tidak (atau kurang) terstruktur, seperti dokumen *word*, *PDF*, kutipan teks dan lain-lain. Sedangkan masukan untuk *Data Mining* adalah data yang terstruktur [3].

Analisis sentimen yang merupakan bagian dari *opinion mining* [4]. Analisis sentimen adalah proses perbandingan untuk mengungkap dan mengkategorikan pendapat yang diungkapkan dalam teks, untuk menentukan apakah tanggapan terhadap topik tertentu positif, negatif, atau netral [5]. Analisis sentimen dapat sebagai solusi dari masalah yang sudah dipaparkan sebelumnya [6]. Sentimen terutama mengacu pada perasaan, emosi, pendapat atau sikap. Dengan meningkatnya *World Wide Web*, orang sering mengungkapkan perasaan mereka melalui internet melalui media sosial, *blog*, *rating* dan *reviews*. Karena ini meningkatkan data tekstual, ada kebutuhan untuk menganalisa konsep mengekspresikan sentimen dan menghitung wawasan untuk menjelajahi bisnis [7].

Ada beberapa pendekatan yang berbeda yang dapat digunakan untuk permasalahan analisis sentimen, seperti *Naive Bayes* [9][10], dan *Bayesian Networks* [10][11], *K-Nearest Neighbor* [12][13], *Maximum Entropy* [14][15], *Neural Network* [16], [17], *Support Vector Machines* [6][18][19][20] dan lain-lain.

Pada penelitian ini algoritma yang akan digunakan adalah *Support Vector Machine* (SVM) dan *Naive Bayes* (NB). Kedua metode ini sering dijadikan metode

dalam penelitian yang membahas sentiment analisis. Pemilihan metode klasifikasi SVM karena memiliki kemampuan generalisasi dalam mengklasifikasikan suatu *pattern*, tidak termasuk data yang dipakai dalam fase pembelajaran metode tersebut [21]. Algoritma SVM digunakan pada analisis sentimen dikarenakan algoritma ini bertujuan sebagai metode klasifikasi kedalam kategori positif dan negatif [6].

Sedangkan Pengklasifikasi *Naive Bayes* (NB) sangat sederhana dan efisien [22]. Selain itu, pengklasifikasi *Naive Bayes* adalah teknik *machine learning* yang populer untuk klasifikasi teks, dan memiliki performa yang baik pada banyak domain [23]. Pada penelitian sebelumnya metode *Naive Bayes* juga digunakan dalam Analisis Sentimen berinternet pada media sosial AMIK BSI Tegal di mana metode *Naive Bayes* digunakan untuk menentukan sentimen publik terhadap AMIK BSI Tegal yang disampaikan di *Instagram*, sehingga membantu usaha untuk melakukan riset *marketing* atas *public opinion* [9].

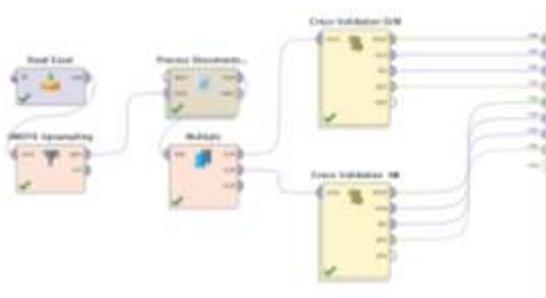
Oleh karena itu dalam penelitian ini digunakan metode SVM dan NB yang diharapkan sesuai dengan permasalahan. Dengan dilakukannya analisis sentimen ini, diharapkan dapat diketahui tanggapan masyarakat terhadap kepuasan kinerja

Kementerian Kesehatan, sehingga dapat dijadikan sarana dan masukan untuk membuat pertimbangan kebijakan Kementerian Kesehatan selanjutnya dalam penanggulangan penyebaran *Covid-19* di Indonesia.

II. METODE PENELITIAN

Penelitian ini dimulai dengan melakukan analisa dan menentukan *query*, Penentuan *query* pada *Twitter* akan membantu dalam pengumpulan data *tweet* sehingga akan lebih terarah. *Query* yang dipilih dalam penelitian ini yaitu Menkes Terawan.

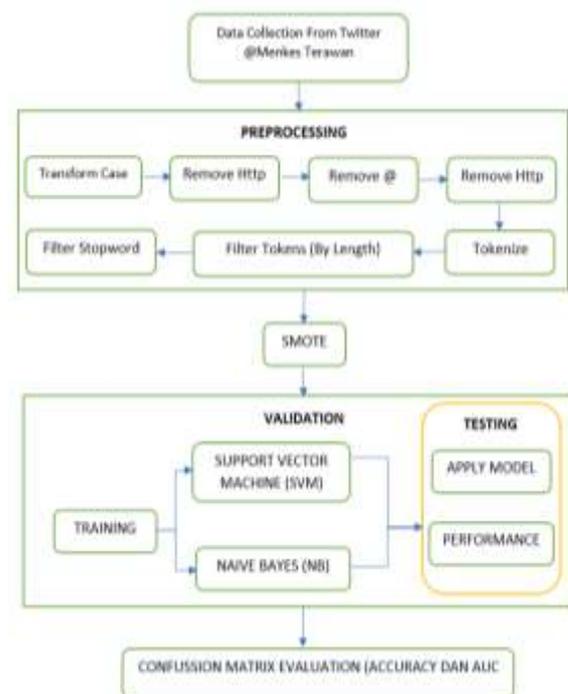
Data *Tweet* diambil dengan menggunakan metode *Crawling* dari sosial media *Twitter* menggunakan *Twitter API* menggunakan *Tools Rapidminer Studio* versi 9.6. Data yang diambil hanya *tweet* berbahasa Indonesia kemudian disimpan ke *Microsoft Excel* dimana setiap *query* terdiri dari 2.000 data. Model analisis data yang dibuat sebagaimana disajikan pada gambar 1.



Gambar 1. Desain Model Komparasi Support Vector Machine dan Naive Bayes

2.1. Model Penelitian

Langkah-langkah yang akan dilakukan pada penelitian ini dituangkan dalam sebuah model kerangka kerja penelitian. Tahapan-tahapan pada model kerangka kerja ini akan dijadikan acuan selama proses penelitian. Model kerangka kerja pada penelitian ini disajikan pada Gambar 2.



Gambar 2. Kerangka Kerja Analisis Sentimen

2.2. Teknik Pengumpulan data

Pada penelitian ini menggunakan data sekunder. Data yang digunakan untuk penelitian dapat berupa data primer maupun data sekunder. Data primer merupakan data yang diperoleh langsung dari sumber datanya, sedangkan data

sekunder merupakan data yang didapatkan dari sumber yang sudah ada.

2.3. Preprocessing

Pada tahapan ini dilakukan proses *preprocessing* yang diantaranya *transform case*, *remove http*, *remove annotation*, *remove hashtag*, *tokenize*, *stopword removal*.

a. Transform Case

Transform case bertujuan untuk mengubah semua teks menjadi huruf kecil. Hal ini bertujuan untuk menghindari masalah ketika dilakukan proses *tokenize*.

b. Remove Http

Remove http bertujuan untuk menghilangkan *URL* yang biasanya terdapat pada teks *tweet*.

c. Remove Annotation

Remove Annotation bertujuan untuk menghapus tanda *annotation* yang seringkali terdapat pada teks *tweet*. *Annotation* merupakan salah satu *noise* yang tidak memiliki arti.

d. Remove Hashtag

Remove Hashtag bertujuan untuk menghapus kata kunci *hashtag*.

e. Tokenize

Tokenize adalah proses yang bertujuan untuk memisahkan setiap kata di dalam teks menjadi sebuah entitas tunggal. Pemisahan kata tersebut dilakukan dengan mengidentifikasi karakter spasi atau tanda baca sebagai pemisah antar kata.

f. Stopword Removal

Stopword Removal adalah proses bertujuan menghilangkan kata-kata umum yang dianggap tidak memiliki makna karena kata-kata tersebut dianggap tidak dapat mewakili konten dokumen. Kata-kata umum tersebut diidentifikasi berdasarkan kata-kata yang terdapat dalam sebuah *dictionary*, dalam penelitian ini digunakan file "*stop-word-list-tala.txt*" yang berisi *stopwords* bahasa Indonesia.

2.4. SMOTE

Dalam pengumpulan Data terkait analisis sentimen setelah dilakukan *preprocessing* sangat mungkin terjadi *imbalance* data. *Imbalance* data merupakan kondisi dimana di antara kedua kelas memiliki selisih yang sangat signifikan. Pemodelan dengan algoritma yang tidak memperhatikan ketidak seimbangan data akan didominasi oleh kelas mayor dan tidak memperhatikan kelas minornya. Hal tersebut akan sangat mempengaruhi hasil model yang diperoleh [24].

Untuk mengatasi masalah ini, maka diperlukan *SMOTE* (*Synthetic Minority Oversampling Technique*), dengan melakukan *up-sampling* atau menambah jumlah data kelas minor agar setara dengan kelas mayor dengan membangkitkan data buatan.

2.5. Validation Data

Tahapan validation data dilakukan dalam proses *K-Fold Cross Validation* dimana evaluasinya sebanyak $K=10$. Algoritma yang digunakan untuk melakukan proses evaluasinya yaitu menggunakan *Support Vector Machine* (SVM) dan *Naïve Bayes* (NB).

a. Support Vector Machine

Support Vector Machine (SVM) adalah metode klasifikasi yang bekerja dengan cara mencari *hyperplane*. SVM merupakan metode yang digunakan untuk klasifikasi dua kelas (*binary classification*). Pada perkembangannya, beberapa metode diusulkan agar SVM bisa digunakan untuk klasifikasi *multi-class* dengan cara mengombinasikan beberapa *binary classifier*.

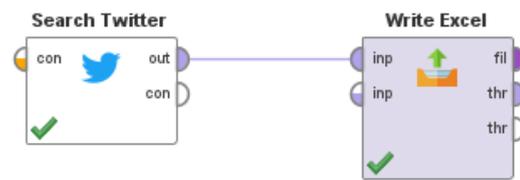
b. Naïve Bayes

Naïve Bayes (NB) adalah algoritma pembelajaran yang didasarkan pada *teorema bayes* dengan menggunakan asumsi yang kuat. *Teorema Bayes* merupakan teori tentang menemukan probabilitas tertinggi sesuatu berdasarkan data yang sudah ada.

III. HASIL DAN PEMBAHASAN

3.1. Pengambilan Dataset

Proses Pengambilan data (*Crawling data*) menggunakan *Tools Rapidminer* diambil dari media sosial *Twitter*.



Gambar 3. Proses *crawling twitter*

Data yang diambil adalah data terbaru dan terpopuler dan hanya *tweet* berbahasa Indonesia dengan jumlah *record* yang di *setting* sebanyak 2.000. Dari 2.000 data tersebut didapatkan 982 data *tweet*, kemudian dilakukan data *cleansing* untuk membersihkan data dari kata yang tidak diperlukan seperti *retweet*, *hashtag* dan untuk menghilangkan 672 duplikasi data yang menghasilkan 310 data *tweet*. Sehingga dataset yang siap diproses untuk *pre-processing* adalah sebanyak 310 baris.

3.2. Pelabelan kelas sentimen

Dari 310 data *tweet* kemudian ditentukan *class attribute*, *class attribute* yang akan digunakan dalam penelitian ini hanya menggunakan 2, yaitu positif dan negatif, karena pada Penelitian ini bertujuan menyimpulkan sistem analisis sentimen ini dapat mengenali kelemahan (berdasarkan komentar negatif) dan kelebihan (berdasarkan komentar positif) secara otomatis dengan baik.

3.3. Preprocessing Data

Data yang telah diperoleh baru akan bisa digunakan setelah melalui beberapa tahap sebelum proses pengujian. tahap ini

adalah tahap *Preprocessing* Data. Data akan diproses melalui beberapa tahap Berikut :



Gambar 4. Proses Preprocessing Data

Dalam Proses *document form Data* terdapat operator *Transform Case*, *remove HTTP*, *remove Annotation(@)*, *remove Hastag(#)*, *Tokenize*, *Filter Token By Length*, *Filter Stopword*.

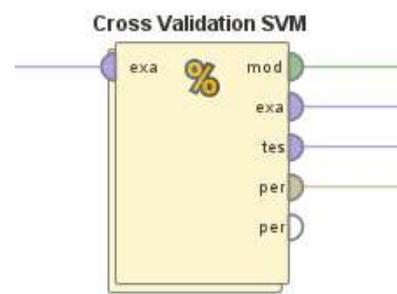
3.4. Pengujian Data

Dalam pengumpulan Data terkait analisis sentimen setelah dilakukan *preprocessing* sangat mungkin terjadi *imbalance* data. *Imbalance* data merupakan kondisi dimana di antara kedua kelas memiliki selisih yang sangat signifikan. Pemodelan dengan algoritma yang tidak memperhatikan ketidakseimbangan data akan didominasi oleh kelas mayor dan tidak memperhatikan kelas minornya. Proses menyeimbangkan jumlah kelas dari *dataset* menggunakan Metode *SMOTE*.

Proses Pengujian ini menggunakan menggunakan algoritma SVM dan NB untuk melakukan analisis sentimen dataset

akan dilakukan pengujian dengan menggunakan metode *Fold Cross Validation*. *Fold cross validation* merupakan metode statistik yang digunakan untuk mengevaluasi kinerja model atau algoritma yang akan memisahkan data menjadi dua *subset* yaitu sebagai data proses pembelajaran (data *training*) dan data validasi (data *testing*) sebagaimana disajikan pada Gambar 5.

Dengan metode 10 *Fold Cross validation*, dataset dibagi menjadi 10 area, dengan masing-masing aspek memberikan informasi yang sama persentase setiap jenis data. 9/10 area data digunakan dalam proses *Training* sehingga membentuk model, sedangkan yang 1/10 area digunakan dalam proses *Testing*. *Training* untuk menghasilkan model dan *testing performance*. sebagaimana disajikan pada Gambar 6 dan 7.



Gambar 5. K-Folds Cross Validation



Gambar 6. Training dan Testing Support Vector Machine



Gambar 7. Training dan Testing Naive Bayes

selanjutnya yaitu melakukan tahap evaluasi atas hasil pengujian data. Evaluasi dilakukan dengan menggunakan nilai akurasi pada tabel *confusion matrix* dan nilai *AUC* pada kurva *ROC*

3.5. Evaluasi Hasil

Setelah melalui proses tahapan pengujian data, berikut ini merupakan evaluasi hasil dari masing-masing algoritma, yaitu *Support Vector Machine* dan *Naive Bayes*:

a. Support Vector Machine

Confusion matrix yang dihasilkan dari algoritma SVM dengan Model *SMOTE* sebagaimana disajikan pada tabel 1.

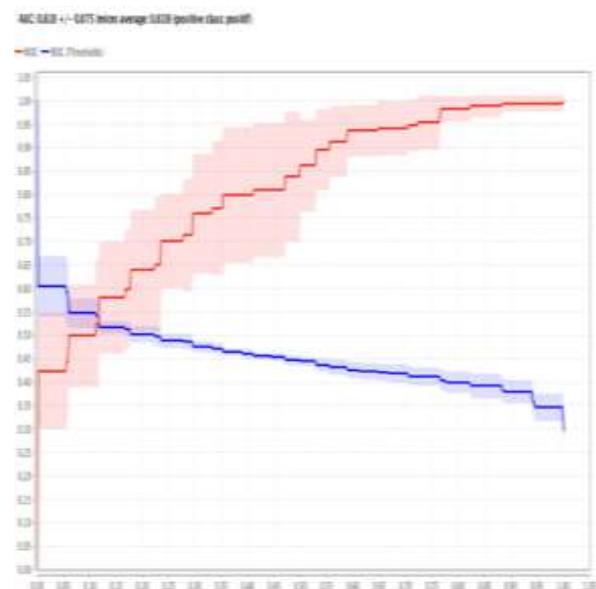
Tabel 1. Confusion Matrix Support Vector Machine

accuracy: 72,57% +/- 5,12% (micro average: 72,54%)			
	<i>true Negatif</i>	<i>true Positif</i>	<i>class precision</i>
pred. Negatif	145	67	68,40%
pred. Positif	28	106	76,10%
<i>class recall</i>	83,82%	61,27%	

Pada tabel 1 tersebut menunjukkan bahwa nilai akurasi yang didapatkan sebesar 72,57% dengan toleransi kesalahan sebesar 5,12%, dengan nilai *true positif* 106 *records* dan *true negatif*

145 *records*. Hal ini menunjukkan bahwa data yang diklasifikasikan sesuai dengan sentimen positif sebanyak 106 *records* dan data yang diklasifikasikan sesuai dengan sentimen negatif sebanyak 145 *records*. Sedangkan data yang seharusnya sentimen positif namun diklasifikasikan negatif (*false negatif*) sebanyak 28 *records* dan data yang seharusnya diklasifikasikan sentimen negatif namun diklasifikasikan positif (*false positif*) sebanyak 67 *records*.

Berdasarkan hasil pengujian *performance* menghasilkan Kurva *Receiver Operating Characteristic* (ROC) sebagaimana pada Gambar 8 dan nilai *Area Under Curve* (AUC) yang didapatkan adalah sebesar 0,810 (*good classification*).



Gambar 8. Grafik ROC Support Vector Machine

b. Naïve Bayes

Confusion matrix yang dihasilkan dari algoritma NB dengan model *SMOTE*. Sebagaimana disajikan pada tabel 2.

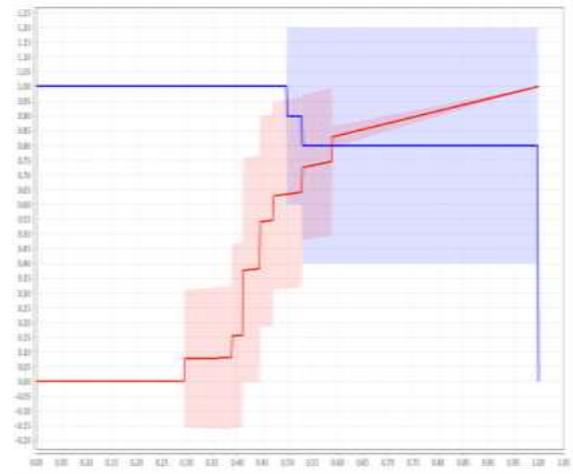
Tabel 2. Confusion Matrix Naïve Bayes

accuracy: 66,45% +/- 3,66% (micro average: 66,47%)			
	<i>true Negatif</i>	<i>true Positif</i>	<i>class precision</i>
pred. Negatif	95	38	71,43%
pred. Positif	78	135	63,38%
<i>class recall</i>	54,91%	78,03%	

Pada Tabel 2 tersebut menunjukkan bahwa nilai akurasi yang didapatkan sebesar 66,45% dengan toleransi kesalahan sebesar 3,66%, dengan nilai *true positif* 135 *records* dan *true negatif* 95 *records*. Hal ini menunjukkan bahwa data yang diklasifikasikan sesuai dengan sentimen positif sebanyak 135 *records* dan data yang diklasifikasikan sesuai dengan sentimen negatif sebanyak 95 *records*. Sedangkan data yang seharusnya sentimen positif namun diklasifikasikan negatif (*false negatif*) sebanyak 78 *records* dan data yang seharusnya diklasifikasikan sentimen negatif namun diklasifikasikan positif (*false positif*) sebanyak 38 *records*.

Berdasarkan hasil pengujian *performance* menghasilkan Kurva *Receiver Operating Characteristic* (ROC) sebagaimana pada Gambar 9 dan nilai *Area Under Curve* (AUC) yang

didapatkan adalah sebesar 0,499 (*poor classification*).



Gambar 9. Grafik ROC Naive Bayes

Berdasarkan analisis rekapitulasi tersebut menunjukkan bahwa evaluasi menggunakan algoritma SVM mendapatkan nilai *accuracy* tertinggi dibandingkan dengan algoritma NB. Berdasarkan hasil penelitian tersebut metode SVM mengungguli metode NB, Hasil penelitian menunjukkan nilai akurasi untuk algoritma NB sebesar 66,45%, Sedangkan nilai akurasi algoritma SVM lebih besar dari NB yaitu sebesar 72,57%. Penelitian ini berhasil mendapatkan algoritma SVM sebagai algoritma yang efektif dan terbaik dalam mengklasifikasikan komentar positif dan komentar negatif terkait dengan sentimen analisis terhadap Kementerian kesehatan.

IV. KESIMPULAN

Dari penelitian yang telah dilakukan maka dapat diambil kesimpulan bahwa

analisis sentimen dapat digunakan untuk mengukur sentimen masyarakat khususnya media sosial *twitter* terhadap Menteri kesehatan. Tujuannya penelitian ini adalah membantu masyarakat menentukan berapa banyak sentimen yang terdapat pada *tweet* opini Bahasa Indonesia yang ada di *Twitter* yang ditujukan kepada Kementerian Kesehatan. Dalam penelitian mengenai analisis sentimen terhadap Menteri Kesehatan diketahui algoritma *Support Vector Machine* memiliki akurasi lebih tinggi mengungguli algoritma *Naïve Bayes*, besar harapan hasil penelitian ini dapat bermanfaat dan membantu pemerintah dalam mendukung pengambilan keputusan selanjutnya dalam menentukan program Pemerintah dalam memutus rantai Covid-19. Selain kesimpulan di atas, penelitian ini juga berupaya memberikan saran untuk penelitian lebih lanjut, terkait dengan jumlah data, penelitian berikutnya dapat meningkatkan jumlah dataset sebagai penelitian karena semakin banyak jumlah datasetnya semakin baik hasilnya melakukan pengujian dengan menggunakan algoritma yang berbeda dari penelitian ini.

DAFTAR PUSTAKA

- [1] M. Kamyab, R. Tao, M. H. Mohammadi, and A. Rasool, "Sentiment analysis on Twitter: A text mining approach to the Afghanistan status reviews," *ACM Int. Conf. Proceeding Ser.*, pp. 14–19, 2018, doi: 10.1145/3293663.3293687.
- [2] J. Weng, E.-P. Lim, J. Jiang, and Z. Qi, *Twitterrank: Finding Topic-Sensitive Influential Twitterers*. 2010.
- [3] J. GoldeNBERG and M. Fresko, "Mine Your Own Business : Market-Structure Surveillance Through Text Mining," vol. 31, no. 3, pp. 521–543, 2012.
- [4] B. Liu, "Sentiment Analysis and Subjectivity," pp. 1–38, 2010, doi: 10.1016/j.dss.2012.05.022.
- [5] R. Jannati, R. Mahendra, C. W. Wardhana, and M. Adriani, "Stance Classification Towards Political Figures on Blog Writing Stance Classification towards Political Figures on Blog Writing," no. February 2019, 2018, doi: 10.1109/IALP.2018.8629144.
- [6] B. Pratama, D. D. Saputra, D. Novianti, and E. P. Purnamasari, "Sentiment Analysis of the Indonesian Police Mobile Brigade Corps Based on Twitter Posts Using the SVM And NB Methods Sentiment Analysis of the

- Indonesian Police Mobile Brigade Corps Based on Twitter Posts Using the SVM And NB Methods,” 2019, doi: 10.1088/1742-6596/1201/1/012038.
- [7] A. Tripathy, A. Agrawal, and S. K. Rath, “Classification of Sentimental Reviews Using Machine Learning Techniques,” *Procedia - Procedia Comput. Sci.*, vol. 57, pp. 821–829, 2015, doi: 10.1016/j.procs.2015.07.523.
- [8] S. Kurniawan, W. Gata, D. A. Puspitawati, M. Tabrani, and K. Novel, “Perbandingan Metode Klasifikasi Analisis Sentimen Tokoh Politik Pada,” vol. 1, no. 10, pp. 2–8, 2021.
- [9] A. Fauzi, A. N. Rais, M. F. Akbar, and W. Gata, “ANALISIS SENTIMEN BERINTERNET PADA MEDIA SOSIAL NAIVE BAYES,” pp. 46–54, 2018.
- [10] N. Friedman, M. Linial, and D. Pe, “Using Bayesian Networks to Analyze Expression Data,” pp. 127–135, 2000, doi: DOI:10.1089/10665270075005096 1.
- [11] I. Chaturvedi, E. Ragusa, P. Gastaldo, R. Zunino, and E. Cambria, “Bayesian Network based Extreme Learning Machine for Subjectivity Detection,” *J. Franklin Inst.*, 2017, doi: 10.1016/j.jfranklin.2017.06.007.
- [12] C. Series, “K-nearest neighbor analysis to predict the accuracy of product delivery using administration of raw material model in the cosmetic industry (PT K-nearest neighbor analysis to predict the accuracy of product delivery using administration of raw material model in the cosmetic industry (PT Cedefindo),” 2019, doi: 10.1088/1742-6596/1367/1/012008.
- [13] L. Dey, “Sentiment Analysis of Review Datasets using Naïve Bayes ’ and K -NN Classifier,” vol. 8, no. 4, p. pp.54-62, 2016, doi: 10.5815/ijieeb.2016.04.07.
- [14] N. D.; Putranti and E. Winarko, “Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine,” vol. 8, no. 1, pp. 91–100, 2014.
- [15] G. Mann, R. Mcdonald, and N. Silberman, “Efficient Large-Scale Distributed Training of Conditional Maximum Entropy Models,” pp. 1–9, 2009.
- [16] A. Severyn, “Twitter Sentiment Analysis with Deep Convolutional

- Neural Networks,” pp. 959–962, 2015, doi: 10.1007/978-3-319-19644-2_60.
- [17] M. G. Santos, Cícero Nogueira dos, “Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts,” pp. 69–78, 2014.
- [18] H. A. Putranto, O. Setyawati, and A. L. Belakang, “Pengaruh Phrase Detection dengan POS -Tagger terhadap Akurasi Klasifikasi Sentimen menggunakan SVM,” vol. 5, no. 4, pp. 252–259, 2016.
- [19] J. Watori, A. Bayhaqy, and H. Pratama, “Sistem Rekomendasi Untuk Pengguna E-Commerce,” vol. 2018, no. Sentika, pp. 23–24, 2018.
- [20] W. Gata and A. Bayhaqy, “Analysis sentiment about islamophobia when Christchurch attack on social media,” *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 18, no. 4, pp. 1819–1827, 2020, doi: 10.12928/TELKOMNIKA.V18I4.14179.
- [21] A. S. Nugroho, A. B. Witarto, and D. Handoko, “Support Vector Machine,” 2003.
- [22] J. Chen, H. Huang, S. Tian, and Y. Qu, “Feature Selection for Text Classification with Naive Bayes,” *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5432–5435, Apr. 2009, doi: 10.1016/j.eswa.2008.06.054.
- [23] Q. Ye, Z. Zhang, and R. Law, “Expert Systems with Applications Sentiment classification of online reviews to travel destinations by supervised machine learning approaches,” *Expert Syst. Appl.*, vol. 36, no. 3, pp. 6527–6535, 2009, doi: 10.1016/j.eswa.2008.07.035.
- [24] Alberto Fernandez, Salvador Garcia, Francisco Herrera, and Nitesh V. Chawla, “SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary,” *J. Artif. Intell. Res.*, vol. 61, pp. 863–905, 2018.
- [25] Sabita, H., & Herwanto, R. (2020). Pantauan Prediktif Covid-19 Dengan Menggunakan Metode SIR dan Model Statistik Di Indonesia. *TEKNIKA*, 14(2), 145-150.

ANALISA DAN PREDIKSI IKLAN LOWONGAN KERJA PALSU DENGAN METODE NATURAL LANGUAGE PROGRAMING DAN MACHINE LEARNING

Hary Sabita¹, Fitria², Riko Herwanto³

¹²³Fakultas Ilmu Komputer, Informatics & Business Institute Darmajaya

Jl. 2.A. Pagar Alam No. 93, Bandar Lampung - Indonesia 35142

Telp. (0721) 787214 Fax. (0721) 700261

e-mail: hary.sabita@darmajaya.ac.id, fitria@darmajaya.ac.id, rikoherwanto@darmajaya.ac.id

ABSTRACT

This research was conducted using the data provided by Kaggle. This data contains features that describe job vacancies. This study used location-based data in the US, which covered 60% of all data. Job vacancies that are posted are categorized as real or fake. This research was conducted by following five stages, namely: defining the problem, collecting data, cleaning data (exploration and pre-processing) and modeling. The evaluation and validation models use Naïve Bayes as a baseline model and Small Group Discussion as end model. For the Naïve Bayes model, an accuracy value of 0.971 and an F1-score of 0.743 is obtained. While the Stochastic Gradient Descent obtained an accuracy value of 0.977 and an F1-score of 0.81. These final results indicate that SGD performs slightly better than Naïve Bayes.

Keywords—NLP, Machine Learning, Naïve Bayes, SGD, Fake Jobs

ABSTRAK

Selama masa pandemi Covid 19, penipuan lowongan pekerjaan atau ketenaga kerjaan meningkat sehingga tekanan ekonomi dan dampak virus Corona telah secara signifikan mengurangi ketersediaan pekerjaan dan kehilangan pekerjaan bagi banyak orang. Kasus seperti ini menghadirkan peluang yang tepat bagi penipu. Kebanyakan *scammer* melakukan ini untuk mendapatkan informasi pribadi dari orang yang mereka *scam*. Penelitian ini menggunakan sample data dari lokasi yang berbasis di AS, yang mencakup 60% dari seluruh data. Lowongan pekerjaan yang diposting, dikategorikan sebagai nyata atau palsu. Penelitian ini dilakukan dengan mengikuti lima tahap, yaitu: pendefinisian masalah, pengumpulan data, pembersihan data (eksplorasi dan pra-pemrosesan) serta pemodelan. Model evaluasi dan validasi menggunakan *Naïve Bayes* sebagai model *baseline* dan *Stochastic Gradient Descent* sebagai model akhir. Untuk model *Naïve Bayes* didapatkan nilai akurasi sebesar 0,971 dan skor F1 sebesar 0,743. Sementara pada *Stochastic Gradient Descent* didapatkan nilai akurasi sebesar 0,977 dan skor F1 sebesar 0,81. Hasil akhir ini menunjukkan bahwa *Stochastic Gradient Descent* memiliki kinerja yang sedikit lebih baik dibandingkan *Naïve Bayes*.

Kata Kunci — NLP, Machines Learning, Naïve Bayes, Stochastic Gradient Descent, Lowongan Pekerjaan Palsu

I. PENDAHULUAN

Selama masa pandemi Covid 19, penipuan lowongan pekerjaan atau ketenagakerjaan meningkat. Menurut *Consumer News and Business Channel* (CNBC), jumlah penipuan ketenagakerjaan meningkat dua kali lipat dibandingkan dengan 2017 [1]. Situasi pasar saat ini telah menyebabkan tingginya pengangguran. Tekanan ekonomi dan dampak virus Corona secara signifikan mengurangi ketersediaan pekerjaan dan kehilangan pekerjaan bagi banyak orang. Kasus seperti ini menghadirkan peluang yang tepat bagi penipu [2-3]. Kebanyakan *scammer* melakukan ini untuk mendapatkan informasi pribadi dari orang yang mereka *scam* [4]. Informasi pribadi dapat berisi alamat, rekening bank, nomor jaminan sosial, dan lain-lain [5]. Para penipu memberi pencari kerja peluang kerja yang sangat menguntungkan dan kemudian meminta uang sebagai imbalan. Atau mereka membutuhkan investasi dari pencari kerja dengan janji pekerjaan. Ini adalah masalah berbahaya yang dapat diatasi melalui teknik *Machine Learning* dan *Natural Language Processing* (NLP) [6-10].

II. METODE PENELITIAN

a. Algoritma dan Teknik

Berdasarkan analisis data yang diperoleh, pemodelan akhir akan menggunakan data teks maupun numerik. Sebelum melakukan pemodelan data, maka kumpulan data akhir ditentukan. Penelitian ini menggunakan kumpulan data dengan fitur-fitur analisis akhir sebagai berikut:

1. *Telecommuting*
2. Kecurangan (*Fraudulent*)
3. Rasio; rasio pekerjaan palsu dan nyata berdasarkan lokasi.
4. Teks; kombinasi dari judul, lokasi, profil_perusahaan, deskripsi, persyaratan, manfaat, pengalaman_wajib, pendidikan_wajib, industri serta fungsi.

Sementara untuk algoritma dan teknik yang digunakan pada penelitian ini adalah:

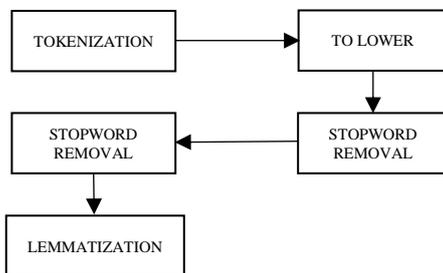
1. NLP
2. Algoritma *Naïve Bayes*
3. Pengklasifikasi *Stochastic Gradient Descent* (SGD)

Proses yang dilakukan adalah, *Naïve Bayes* dan pengklasifikasi SGD dibandingkan untuk melihat akurasi dan skor F1, selanjutnya model akhir ditentukan. Pada penelitian ini, *Naïve Bayes* digunakan untuk menghitung probabilitas kondisional dua peristiwa berdasarkan probabilitas kemunculan setiap peristiwa. Pengklasifikasian SGD

mengimplementasikan rutinitas pembelajaran penurunan gradien stokastik sederhana yang mendukung fungsi kerugian dan pinalti berbeda untuk klasifikasi. Model ini digunakan pada teks dan numerik secara terpisah, kemudian hasil akhir digabungkan.

b. Metodologi

Untuk pemrosesan data, langkah – langkah yang dilakukan adalah:

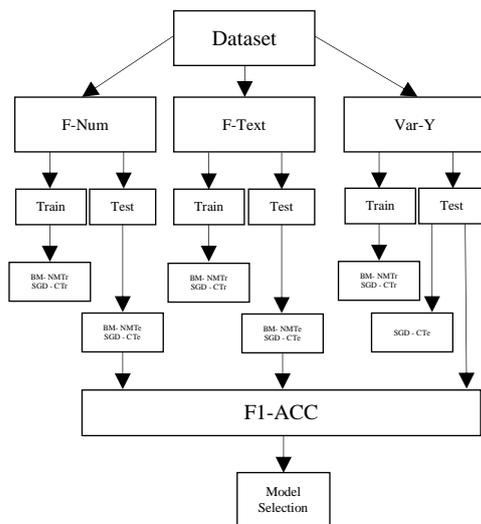


Gambar 1. II. Pemrosesan Data

- *Tokenization*: Data tekstual dibagi menjadi beberapa unit yang lebih kecil. Data dipecah – pecah menjadi kata – kata.
- *To Lower*: Kata – kata yang dipisahkan, diubah menjadi huruf kecil.
- *Stopword Removal*: Adalah kata – kata yang tidak menambahkan banyak arti pada kalimat, seperti: the, a, an, he, have dan lain – lain.
- *Lemmatization*: Proses pengelompokan lemmatisasi dimana bentuk kata – kata yang berubah digunakan bersama – sama.

c. Implementasi

Dataset dibagi menjadi teks, numerik dan variabel y. Dataset teks diubah menjadi matriks frekuensi *term* untuk analisis lebih lanjut. Dengan menggunakan *free software*, dan memungkinkan kita melakukan beragam pekerjaan dalam *Data Science*, seperti regresi, klasifikasi, pengelompokan, *data preprocessing*, *dimensionality reduction* dan *model selection* (perbandingan, validasi, dan pemilihan parameter maupun model) atau *Scikit learning*, dataset tersebut dipecah menjadi dataset *test* dan *training*. Model dasar *Naïve Bayes* dan model SGD dilatih menggunakan set *train*, sekitar 70% dari kumpulan data. Hasil akhir dari model, berdasarkan dua set pengujian – numerik dan teks, digabungkan sehingga jika kedua model mengatakan bahwa titik data tertentu bukan hanya penipuan, maka lowongan kerja tersebut adalah palsu. Hal ini dilakukan untuk mengurangi bias algoritma *Machine Learning* terhadap kelas mayoritas. Model yang dilatih digunakan pada set pengujian, dengan tujuan untuk mengevaluasi performa model. Akurasi dan skor F1 dari *Naïve Bayes* dan SGD dibandingkan, selanjutnya pemilihan model akhir untuk analisis.



Gambar 2. Proses Pemilihan Model Akhir

Keterangan:

F-Num = Fitur numerik (Telecommuting, Rasio, Jumlah Karakter)

F-Text = Fitur text (idf dan jumlah matriks

Var-Y = Variabel Y (Kecurangan-Fraudulent)

BM-NMTr = *Baseline Model – Naïve Model*

Training

SGD-CTr = *Stochastic Gradient Descent Classifier*

Training

BM-NMTe = *Baseline Model – Naïve Model Test*

SGD-CTe = *Stochastic Gradient Descent Classifier*

Test

F1-ACC = F1 and Accuracy Calculation and Comparison

d. Metrik

Penelitian ini melakukan evaluasi

dengan menggunakan dua metrik, yaitu:

1. Akurasi

Rumus metrik ini didefinisikan sebagai berikut:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

Keterangan:

TP = *True Positive*

TN = *True Negative*

FP = *False Positive*

FN = *False Negative*

Rumus ini menghasilkan rasio dari semua poin data yang dikategorikan dengan benar ke semua poin data. Hal ini sangat

berguna untuk mengidentifikasi lowongan pekerjaan nyata atau palsu. Kelemahan pada metrik ini adalah, algoritma *Machine Learning* cenderung menyukai kelas yang dominan. Data yang ada pada penelitian ini kelasnya tidak seimbang, akurasi yang tinggi akan menunjukkan seberapa baik model kita kategorikan sebagai kategori negatif (lowongan pekerjaan nyata).

2. Skor F1

Ini adalah ukuran akurasi model pada kumpulan data. Dijelaskan pada rumus (2):

$$Skor F1 = \frac{TP}{TP + \frac{FP+FN}{2}} \quad (2)$$

Keterangan:

TP = *True Positive*

FP = *False Positive*

FN = *False Negative*

Skor F1 digunakan, karena dalam skenario ini, negatif palsu dan positif palsu sangat penting. Model ini diperlukan untuk mengidentifikasi kedua kategori dengan skor setinggi mungkin, karena biaya keduanya sangat tinggi.

III. HASIL DAN PEMBAHASAN

a. Eksplorasi Data

Pada penelitian ini, dataset yang digunakan berjumlah 17.880 data dan 18 fitur. Tabel 1, menerangkan fitur – fitur dataset yang digunakan, serta *script* informasi untuk data fitur.

```
In [1]: import pandas as pd
In [4]: lowonganPalsu = pd.read_csv('fake_job_postings.csv')
In [7]: lowonganPalsu.columns
Out[7]: Index(['job_id', 'title', 'location', 'department', 'salary_range',
            'company_profile', 'description', 'requirements', 'benefits',
            'telecommuting', 'has_company_logo', 'has_questions', 'employment_type',
            'required_experience', 'required_education', 'industry', 'function',
            'fraudulent'],
            dtype='object')
```

Tabel 1. Fitur-Fitur Dataset yang Digunakan

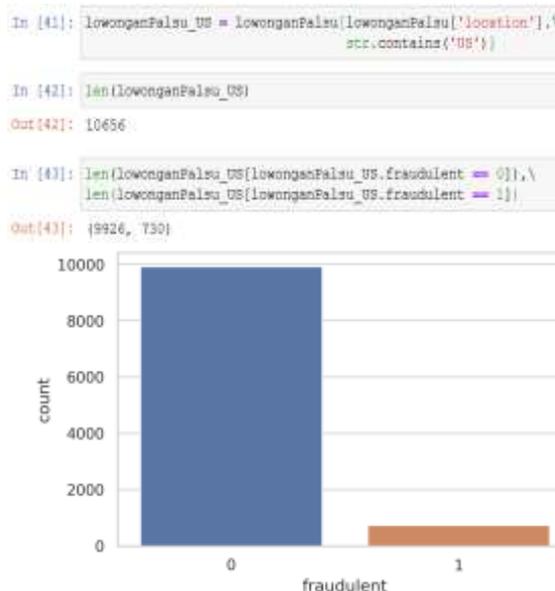
No	Variabel	Tipe data	Keterangan
1	job_id	int	nomor identifikasi lowongan
2	title	text	nama jabatan atau pekerjaan
3	location	text	lokasi pekerjaan
4	department	text	departemen pekerjaan
5	salary_range	text	gaji yang ditawarkan
6	company_profile	text	informasi perusahaan
7	description	text	ringkasan perusahaan
8	requirements	text	informasi kebutuhan awal pekerjaan
9	benefits	text	keuntungan yang ditawarkan
10	telecommuting	boolean	informasi pekerjaan remote atau on site
11	has_company_logo	boolean	apakah perusahaan memiliki logo
12	has_questions	boolean	apakah pekerjaan memiliki pertanyaan
13	employment type	text	kategori pekerjaan
14	required_experience	text	kebutuhan pengalaman pekerjaan
15	required_education	text	kebutuhan jenjang pendidikan
16	industry	text	jenis industri
17	function	text	fungsi pekerjaan secara umum
18	fraudulent	boolean	0; nyata, 1; palsu

Karena sebagian besar tipe data ini adalah teks, maka tidak digunakan statistik ringkasan. Tipe data *integer* hanya ada pada *job_id*, dan tidak relevan untuk analisis ini. Selanjutnya dataset dieksplorasi untuk mengidentifikasi nilai *null*. Variabel yang banyak memiliki nilai yang hilang adalah *department* dan *salary_range*. Untuk nilai *salary_range* yang hilang, diisi dengan nilai rata – rata dari kolom tersebut. Berikut *script* yang informasinya.

```
In [11]: lowonganPalsu.isnull().sum()
Out[11]: job_id          0
         title           0
         location       346
         department    11547
         salary_range   15012
         company_profile 3308
         description     1
         requirements   2695
         benefits       7210
         telecommuting  0
         has_company_logo 0
         has_questions  0
         employment_type 3471
         required_experience 7050
         required_education 8105
         industry       4903
         function       6455
         fraudulent     0
         dtype: int64
```

Dataset ini berisi lowongan pekerjaan dari beberapa negara, dalam rentan tahun 2020. Banyak data menggunakan bahasa yang berbeda, data yang digunakan adalah data yang berbasis lokasi di AS, yang mencakup 60% dari kumpulan data. Data ini dijadikan sampel karena semua dalam bahasa Inggris, agar mudah ditafsirkan. Selain itu, lokasi dibagi menjadi negara bagian dan kota untuk analisis lebih lanjut. Setelah dilakukan eksplorasi data, maka dataset akhir memiliki sampel sebanyak 10.656. Gambar 3. menjelaskan plot perhitungan untuk lowongan pekerjaan

yang menghasilkan 9926 (93,14%) lowongan pekerjaan nyata dan hanya 730 (6,85%) lowongan pekerjaan palsu, serta *script* untuk jumlah lowongan yang ada di AS.

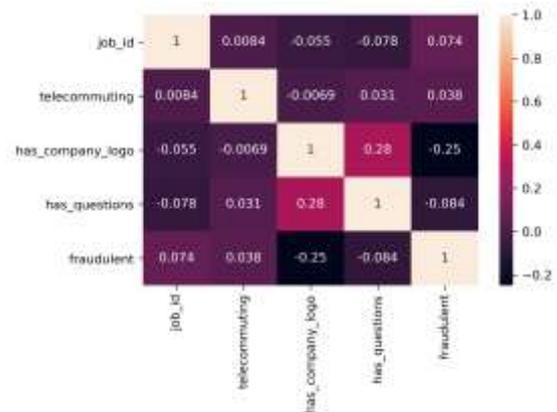


Gambar 3. Jumlah Lowongan Kerja Nyata dan Palsu Di AS

b. Analisis Eksplorasi Data

Langkah pertama yang dilakukan untuk memvisualisasikan dataset pada penelitian ini adalah dengan cara membuat matriks korelasi. Tujuannya adalah untuk mempelajari hubungan data numerik. Matriks korelasi tidak menunjukkan korelasi positif atau negatif yang kuat antara data numerik. Gambar 4. menjelaskan hubungan korelasi matriks pada dataset dan informasi *script* nya.

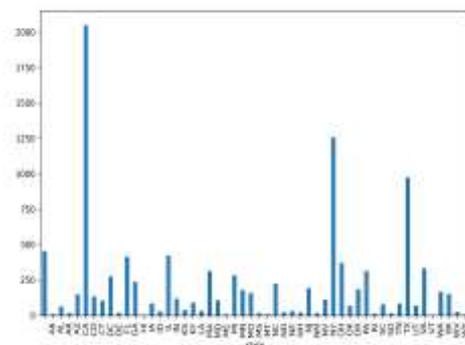
```
In [87]: sns.heatmap(lowonganPalsu_US.corr(), annot=True);
```



Gambar 4. Korelasi Hubungan Matriks pada Dataset

Selanjutnya fitur yang dieksplorasi adalah fitur tekstual. Eksplorasi data ini dimulai dari lokasi yang ada di negara bagian AS. Gambar 5., jumlah pekerjaan berdasarkan lokasi dan informasi *script* nya.

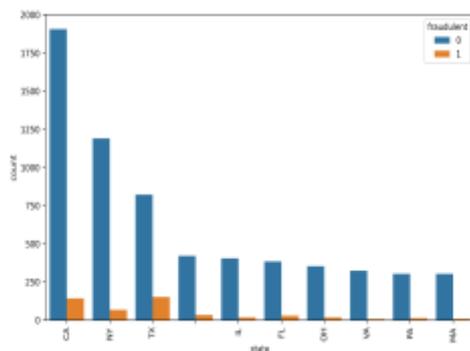
```
In [93]: plt.figure(figsize=(10,6))
      lowonganPalsu_US.groupby('state').fraudulent.count().plot(kind='bar');
      plt.savefig('lokerAS5.png', dpi=300)
      plt.show()
```



Gambar 5. Jumlah Lowongan Pekerjaan Berdasarkan Lokasi

Grafik pada Gambar 5., menunjukkan bahwa negara bagian; California, New York dan Texas memiliki jumlah lowongan pekerjaan tertinggi. Untuk mengeksplorasi data lebih jauh, Gambar 6, menampilkan distribusi jumlah pekerjaan nyata dan palsu di 10 negara bagian teratas beserta informasi *script* nya.

```
In [192]: plt.figure(figsize=(10,6))
sns.countplot(x='state', data=lowonganPalsu_US, hue='fraudulent', \
              order=lowonganPalsu_US['state'].\
              value_counts().iloc[:10].index)
plt.xticks(rotation=90)
plt.savefig('l1berAS6.png', dpi=300)
plt.show()
```



Gambar 6. Distribusi Lowongan Pekerjaan Nyata Dan Palsu Di 10 Negara Bagian

Hasilnya menunjukkan bahwa Texas dan California memiliki kemungkinan lowongan pekerjaan palsu lebih tinggi dibandingkan negara bagian lainnya.

c. Hasil Evaluasi dan Validasi Model

Model terakhir yang digunakan untuk analisis ini adalah SGD. Hal ini dilakukan berdasarkan hasil Pengukuran yang dibandingkan dengan model dasar. Nilai perbandingan akurasi dan skor F1 antara Naïve Bayes (model *baseline*) dan SGD (model akhir), disajikan pada Tabel 2, berikut dengan informasi *script* nya.

```
In [219]: from sklearn.naive_bayes import MultinomialNB
from sklearn import metrics
from sklearn.model_selection import train_test_split

nb_classifier = MultinomialNB()
nb_classifier.fit(x_train, y_train)
pred = nb_classifier.predict(x_test)

In [220]: metrics.accuracy_score(y_test, pred) # Nilai akurasi untuk Naive-Bayes/ model baseline
Out[220]: 0.97115501106605

In [221]: metrics.f1_score(y_test, pred) # Nilai skor F1 untuk Naive Bayes/ model baseline
Out[221]: 0.7435897435897438
```

```
In [222]: from sklearn.linear_model import SGDClassifier
clf_log = SGDClassifier(loss='log').fit(x_train, y_train)
pred_log = clf_log.predict(x_test)

In [223]: metrics.accuracy_score(y_test, pred_log)
Out[223]: 0.977689787383247

In [224]: clf_sgd = SGDClassifier(loss='log').fit(x_train_sgd, y_train)
pred_sgd = clf_sgd.predict(x_test_sgd)
metrics.accuracy_score(y_test, pred_sgd)

Out[224]: 0.9336384479353268

In [227]: prediction_array = []
for i, j in zip(pred_sgd, pred_log):
    if i == 1 and j == 0:
        prediction_array.append(1)
    else:
        prediction_array.append(0)

In [228]: metrics.accuracy_score(y_test, prediction_array) # Nilai akurasi untuk SGD
Out[228]: 0.9788304826135608

In [229]: metrics.f1_score(y_test, prediction_array) # Nilai skor F1 untuk SGD
Out[229]: 0.814643388244881
```

Tabel 2. Perbandingan Nilai Akurasi dan Skor F1

Model	Akurasi	Skor F1
Naïve Bayes	0,971	0,743
SGD	0,977	0,81

IV. SIMPULAN

Berdasarkan nilai akhir yang didapatkan, model SGD memiliki kinerja lebih baik dari model *Naïve Bayes*. Model SGD dapat mengidentifikasi pekerjaan nyata dengan akurasi yang sangat tinggi sebesar 0,977 dan Skor F1 sebesar 0,81.

PENELITIAN LANJUTAN

Dataset yang digunakan dalam penelitian ini sangat tidak seimbang. Kebanyakan lowongan pekerjaannya adalah nyata dan sedikit sekali yang palsu. Karena itu, pekerjaan nyata diidentifikasi dengan cukup baik. Untuk itu, penelitian berikutnya diharapkan dapat menambahkan teknik yang dapat menghasilkan sampel kelas minoritas,

sehingga dataset menjadi seimbang dan memberikan hasil yang lebih baik.

DAFTAR PUSTAKA

- [1] Habiba, S. I. (2021). A Comparative Study on Fake Job Post Prediction Using Different Data Mining Technique. In *2021 2nd International Conference on Robotics, Electrical and Signaling Processing Techniques (ICREST)*, 543-546.
- [2] Ranparia, D., Kumari, S., & Sahani, A. (2020, November). Fake Job Prediction using Sequential Network. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)* (pp. 339-343). IEEE.
- [3] Vidros, S., Koliass, C., Kambourakis, G., & Akoglu, L. (2017). Automatic detection of online recruitment frauds: Characteristics, methods, and a public dataset. *Future Internet*, 9(1), 6.
- [4] Zamir, H. (2020). Cybersecurity and Social Media. *Cybersecurity for Information Professionals: Concepts and Applications*.
- [5] Wilson, J. (2018). Scamming the scammers with their own tricks. *Computer Fraud & Security*, 2018(9), 14-16.
- [6] B Keerthana, A. R. (2021). *Accurate Prediction of Fake Job Offers Using Machine Learning*. Singapore: Springer.
- [7] Ahmed, A. A. A., Aljabouh, A., Donepudi, P. K., & Choi, M. S. (2021). Detecting Fake News using Machine Learning: A Systematic Literature Review. *arXiv preprint arXiv:2102.04458*.
- [8] Ranparia, D., Kumari, S., & Sahani, A. (2020, November). Fake Job Prediction using Sequential Network. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)* (pp. 339-343). IEEE.
- [9] Mahbub, S. &. (2018). Using Contextual Features for Online Recruitment Fraud Detection. *Association For Information Systems*, -.
- [10] Oshikawa, R., Qian, J., & Wang, W. Y. (2018). A survey on natural language processing for fake news detection. *arXiv preprint arXiv:1811.00770*.
- [11] Sabita, H., & Herwanto, R. (2020). Pantauan Prediktif Covid-19 Dengan Menggunakan Metode SIR

- dan Model Statistik Di Indonesia. *TEKNIKA*, 14(2), 145-150.
- [12] Wibowo, H., & Indriyani, F. (2018, October). K-Nearest Neighbor Method For Monitoring Of Production And Preservation Information (Treatment) Of Rubber Tree Plant. In *International Conference on Information Technology and Business (ICITB)* (pp. 29-44).
- [13] Herwanto, R., Purbo, O. W., & Sriyanto, S. (2020). Membandingkan Performa Antara Hyperledger Dan Mysql. *Jurnal Informatika*, 20(1), 89-100.

PREDIKSI TINGKAT KESUKSESAN PROMOSI BANK DENGAN ALGORITMA DNN

Oscar¹, Nurlaelatul Maulidah², Annida Purnamawati³, Destiana Putri⁴, Hilman F. Pardede⁵

¹²³⁴⁵Sekolah Pascasarjana Ilmu Komputer, STMIK Nusa Mandiri, Jakarta, Indonesia

⁵Pusat Penelitian Informatika, Lembaga Ilmu Pengetahuan Indonesia, Indonesia

e-mail : 14002331@nusamandiri.ac.id, 14002433@nusamandiri.ac.id,

14002403@nusamandiri.ac.id, 14002366@nusamandiri.ac.id, hilman@nusamandiri.ac.id

ABSTRACT

Telemarketing is one effective way for promoting products. However, it is often difficult to measure the success of telemarketing. Therefore, a way to predict the success rate of telemarketing, and hence strategies could be planned to increase the success rate. In this study, we evaluate several implementations of machine learning for prediction the success of telemarketing. The evaluated methods are Deep Neural Network (DNN), Random Forest, and K-nearest neighbor (K-NN). We validate our experiments using 10-fold cross validation and our experiments show that DNN with 3 hidden layers outperforms other methods. Accuracy of 90% is achieved with the DNN. It is better than Random Forest and KNN that achieve accuracies of algorithm and 88% and 89%.

Keywords— *Bank Marketing, DNN, KNN, Random Forest.*

ABSTRAK

Telemarketing adalah promosi yang efektif dalam mempromosikan produk. Strategi pemasaran ini dilakukan bank dengan menawarkan produk kepada nasabah, salah satu produk yang ditawarkan bank adalah deposito. Sulitnya mengetahui keputusan klien Telemarketing untuk menempatkan deposito di bank menyebabkan bank selalu menghadapi ancaman krisis keuangan. Dalam mendukung telemarketing untuk meningkatkan tingkat keberhasilan, keputusan klien untuk memprediksi keberhasilan dibuat. Dalam penelitian ini dilakukan percobaan dan pengujian dengan beberapa metode yaitu Klasifikasi yang meliputi Deep Neural Network (DNN), Random Forest, K-nearest neighbour (K-nn). Dengan menambahkan teknik data mining berupa Cross Validation yang bertujuan untuk mendapatkan hasil akurasi yang maksimal. Metode ini sering disebut dengan k-fold cross validation, dimana percobaan k kali untuk satu model dengan parameter yang sama dan menggunakan aplikasi Multi Layer Perceptron (MLP) hasil investigasi dalam pengujian prediksi tingkat keberhasilan promosi bank menggunakan 20: 80 data pengujian dan data pelatihan. Dengan menggunakan arsitektur yang baik berdasarkan hasil pengujian yaitu 3 hidden layer, learning rate dan 10 epoch testing experiment diperoleh nilai akurasi terbaik 90% dengan metode DNN dan akurasi 89% dengan algoritma KNN dan 88% dengan algoritma Random Forest.

Kata Kunci— Bank Marketing, DNN, KNN, Random Forest.

I. PENDAHULUAN

Promosi konvensional produk bank biasanya melalui brosur dan kunjungan karyawan ke nasabah secara langsung. Namun, ini kurang efektif karena eksposur yang terbatas kepada klien potensial. Jadi, strategi telemarketing adalah salah satu cara untuk meningkatkan eksposur produk kepada calon klien. Strategi ini sangat membantu para pemasar dan pelaku usaha untuk memperkenalkan produk dan jasanya kepada masyarakat luas dalam waktu yang singkat [1]. Selain itu, produk yang ditawarkan melalui telemarketing cenderung mudah diterima, karena ditujukan langsung kepada konsumen.

Saat ini, metode telemarketing biasanya dilakukan oleh beberapa perusahaan operator besar untuk menawarkan produknya kepada apa yang mereka anggap sebagai calon klien. Salah satu faktor keberhasilan telemarketing adalah jika produk tersebut dibutuhkan oleh pelanggan. Namun, seringkali sulit untuk mengidentifikasinya. Selain itu, kesuksesan telemarketing seringkali sulit untuk diketahui atau diukur. Dengan mengetahui sukses tidaknya sebuah telemarketing, kita bisa mengidentifikasi pola untuk menemukan calon klien untuk

masa depan. Oleh karena itu, ini dapat digunakan untuk merancang strategi telemarketing [4]. Penambangan data dan pembelajaran mesin dapat digunakan untuk tujuan yang disebutkan di atas [2].

Pada studi sebelumnya oleh [2], algoritma Naïve Bayes, Decision Tree, dan Neural Network digunakan untuk dataset Bank Marketing. Hasil penelitian menunjukkan bahwa pengklasifikasi Decision Tree berkinerja lebih baik daripada pengklasifikasi Naïve Bayes dan Neural Network. Dalam [3], Logistic Regression, Decision Tree, Naïve Bayes dan Random diimplementasikan untuk tujuan yang sama. Random Forest Ensemble menjadi yang terbaik saat data pertama kali diseimbangkan. Studi ini juga menemukan algoritma Decision Tree (CART) ternyata memiliki kinerja yang lebih baik daripada algoritma Logistic Regression (LR) dan Naïve Bayes (NB).

Meskipun metode konvensional dapat mencapai akurasi untuk prediksi keberhasilan dan kegagalan pemasaran jarak jauh, masih dapat ditingkatkan. Salah satu metode potensial adalah DNN. Ini karena, relasi nonlinier antar fitur data dapat dimodelkan dengan lebih baik menggunakan DNN daripada metode

konvensional lainnya. Di sini, kami mengusulkan arsitektur DNN untuk tugas ini dengan memvariasikan jumlah lapisan tersembunyi DNN. Memiliki lebih banyak hidden layers memungkinkan DNN untuk membuat model hubungan nonlinier dengan lebih baik. Namun, kekurangannya adalah waktu komputasi dan data yang lengkap. Eksperimen kami menunjukkan bahwa DNN dengan 3 lapisan tersembunyi lebih baik daripada metode pembelajaran mesin konvensional yang dievaluasi.

II. METODE PENELITIAN

Pembelajaran mesin adalah metode analisis data dan merupakan bagian dari kecerdasan buatan yang didasarkan pada kemampuan untuk mempelajari data yang sudah ada sebelumnya, mengenali pola dan distribusi data serta digunakan untuk membuat keputusan. Sistem ini membangun model analitik otomatis yang meminimalkan campur tangan manusia.

a. Deposito

Deposito berjangka adalah jenis simpanan yang diterbitkan oleh bank, berbeda dengan rekening giro dan tabungan dimana simpanan berjangka mengandung unsur yang memiliki masa jatuh tempo yang lebih panjang (maturity) dan tidak dapat ditarik setiap saat. Bank sepenuhnya didanai oleh simpanan,

persaingan bank yang ketat akan meningkatkan dana bank dengan menaikkan suku bunga simpanan, suku bunga simpanan sebagai ukuran intensitas simpanan dalam persaingan pasar. Dalam hal ini, bank dapat memanfaatkan para deposan dengan menginvestasikan kekayaan asetnya di perusahaan [4].

b. Deep Learning

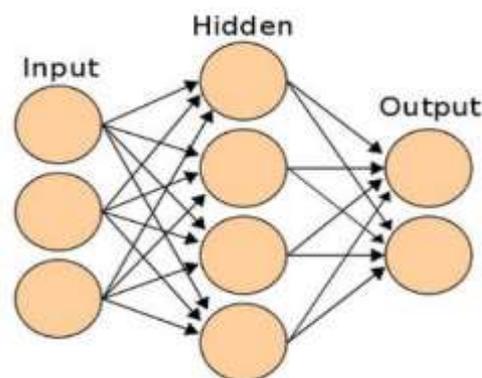
Deep Learning adalah teknik dalam pembelajaran mesin yang memanfaatkan banyak lapisan pemrosesan informasi nonlinier untuk melakukan ekstraksi fitur, pemrosesan pola, dan klasifikasi. Deep Learning merupakan salah satu bagian dari Machine Learning yang memanfaatkan teknik Neural Network atau jaringan syaraf tiruan dalam menyelesaikan suatu masalah [5].

Deep Learning pada dasarnya adalah area baru dalam pembelajaran mesin. Metode ini dilatarbelakangi oleh penelitian kecerdasan buatan yang bertujuan untuk meniru kemampuan manusia dalam mengamati, menganalisis, mempelajari dan mengambil keputusan, terutama untuk memecahkan masalah yang sulit dan kompleks. Pada prinsipnya metode DL dapat dikatakan sebagai salah satu penelitian di bidang pemrosesan sinyal, jaringan saraf tiruan, pemodelan grafis,

optimasi dan pengenalan pola. Reputasi penelitian saat ini di DL didorong oleh peningkatan yang signifikan dalam kekuatan pemrosesan chip komputer, penurunan drastis harga perangkat keras, dan penelitian lanjutan di bidang pembelajaran mesin [6].

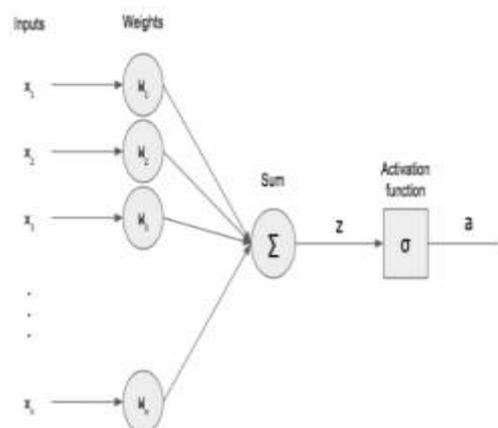
c. Deep Neural Network (DNN)

Algoritma DNN (Deep Neural Networks) merupakan algoritma berbasis neural network yang dapat digunakan untuk pengambilan keputusan. Deep Neural Networks mempunyai tujuan meniru cara kerja otak manusia dengan metode Multi Layer. DNN ini terdiri dari beberapa Lapisan Tersembunyi dengan koneksi antar lapisan tetapi tidak ada koneksi antar unit di setiap lapisan. Pendekatan ini memungkinkan data yang kompleks untuk dimodelkan dengan lebih mudah. Metode ini memiliki arsitektur yang mirip dengan Jaringan Syaraf Tiruan (JST), dengan Pelatihan yang Diawasi. Dengan mengidentifikasi masukan dan mencocokkannya dengan pola yang sulit [7].



Gambar 1. DNN Layer

Deep Neural Network (DNN) adalah tiruan dengan jumlah hidden layers yang lebih besar antara lapisan masukan dan keluaran. Pada studi ini dibuat deep neural network (DNN) dengan multi layer yang saling berhubungan antara input dan output layer, dimana setiap hidden layer memiliki 3 variasi. Arsitektur DNN yang diusulkan menggunakan 2, Encoder dan decoder, variasi Learning 3 dan variasi Optimizer 3 dalam algoritma klasifikasi yang mendasarinya [8].



Gambar 2. Single Perceptron

Pada Gambar 1, menggambarkan bahwa setiap neuron atau satu perceptron

menerima masukan dari lapisan masukan (x_i). Setiap fitur x memiliki bobotnya sendiri (w). Kemudian untuk setiap neuron, lakukan operasi linier dengan menjumlahkan jumlah bobot semua input $\Sigma (w_n x_n)$ dan menambahkan bias (b). Hasil (z) akan dioperasikan dengan fungsi aktivasi (σ).

Jadi secara formal dapat didefinisikan dengan persamaan (1):

$$y_i = \sigma\{\sum_{i=0}^n w^i x^i + b\} \quad (1) \square$$

Fungsi aktivasi mengoperasikan hasil pembobotan input dan bias untuk menentukan apakah suatu neuron diaktifkan atau tidak. Dengan memanipulasi data melalui proses penurunan gradien yang menghasilkan keluaran untuk jaringan saraf [13].

Fungsi aktivasi sigmoid dikaitkan dengan fungsi logistik atau non-linier yang biasa digunakan di FFNN. Fungsi sigmoid pada persamaan (2) menghasilkan nilai mendekati 0 dan 1. Sigmoid diterapkan pada lapisan keluaran.

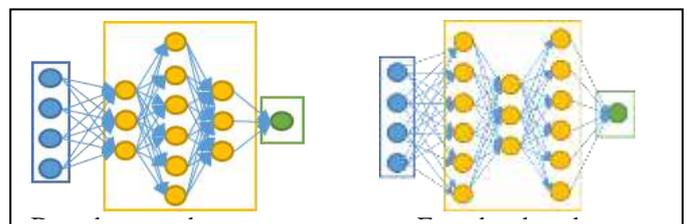
$$\sigma = \frac{1}{(1 + \exp^{-z})} \quad \square(2)$$

Fungsi aktivasi Rectified Linear Unit (ReLU) menghasilkan operasi batas pada

setiap elemen masukan di mana nilai yang kurang dari nol akan menjadi nol, lebih dari nol sebagai fungsi linier. Fungsi ULT dengan persamaan (3) diterapkan ke lapisan tersembunyi di DNN [14].

$$\sigma = \max(0, z) = \begin{cases} z_i, & \text{if } z_i \geq 0 \\ z, & \text{if } z_i < 0 \end{cases} \quad \square(3)$$

d. Arsitektur



Gambar 3. Arsitektur Deep Learning Neural Network

Ada beberapa jenis arsitektur jaringan saraf, termasuk jaringan feedforward, jaringan berulang, dan jaringan fungsi basis radial. Model feed forward neural network (FFNN) merupakan model yang sering digunakan untuk tujuan prediksi [12].

e. Random Forest

Metode Ensemble Random Forest dalam hal pembelajaran digunakan untuk klasifikasi, regresi, dan tugas lainnya. Random Forest acak pertama kali diusulkan oleh Tin Kam Ho dan dikembangkan lebih lanjut oleh Leo Breiman dan Adele Cutler. Performa Random Forest acak diadaptasi dari

Decision Tree, dengan setiap pohon dikembangkan dari sampel bootstrap berdasarkan data pelatihan. Saat mengembangkan pohon, subset atribut diambil secara acak dari atribut terbaik yang akan dipilih untuk dipisahkan. Model akhir dari Random Forest acak didasarkan pada hasil dari seluruh subset pohon yang telah dikembangkan [9].

Random Forest adalah metode pembelajaran melalui pembangunan Decision Tree melalui pelatihan. Ada tiga aspek penting dalam metode random forest, yaitu:

1. Lakukan pengambilan sampel bootstrap untuk membuat Decision Tree.
2. Setiap Decision Tree diprediksi dengan prediktor acak.
3. Kemudian Random Forest acak membuat prediksi dengan menggabungkan hasil dari setiap Decision Tree dengan cara suara terbanyak untuk klasifikasi atau rata-rata untuk regresi [10].

f. K-nearest neighbor

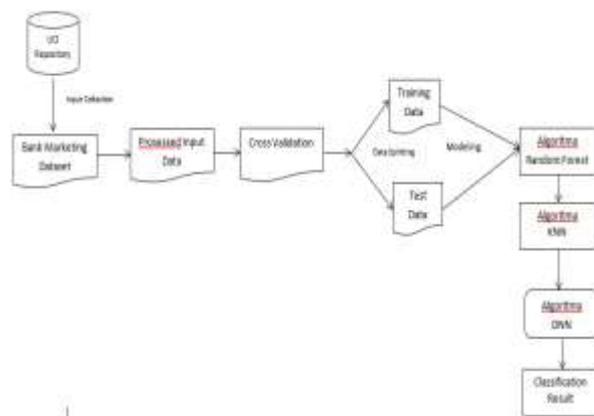
Algoritma K-nearest neighbor (k-NN atau KNN) merupakan metode untuk mengklasifikasikan objek berdasarkan data pembelajaran yang paling dekat dengan

objek, akurasi algoritma k-NN sangat dipengaruhi oleh ada tidaknya fitur yang tidak relevan. , atau jika bobot fitur tidak setara dengan relevansinya dengan klasifikasi [11]. Dalam mengklasifikasikan algoritma K-NN dengan cara mencocokkan bobot dan fitur yang terdapat pada dataset. Penggunaan K dalam algoritma K-NN digunakan untuk menentukan jumlah tetangga yang ada dan digunakan untuk mengambil keputusan [12]. Sebagian besar penelitian tentang algoritma ini membahas tentang bagaimana memilih dan memberi bobot pada fitur, agar performansi klasifikasi lebih baik, KNN juga merupakan contoh dari teknik lazy learning yaitu teknik menunggu query datang sama. sebagai data pelatihan. [11].

III. HASIL DAN PEMBAHASAN

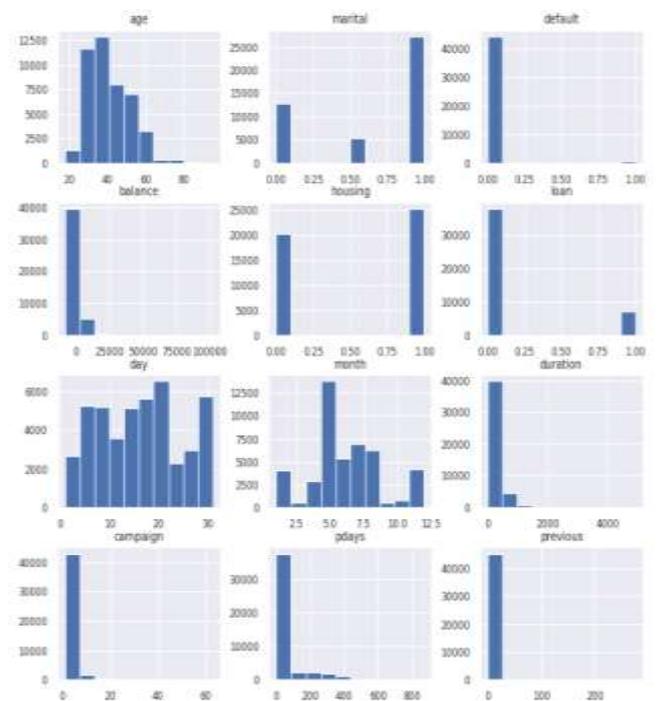
Dalam penelitian ini, data yang digunakan adalah Kumpulan Data Pemasaran Bank yang diperoleh dari data publik yaitu UCI Machine Learning Repository dengan link: <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>.

Dalam penelitian ini akan dilakukan beberapa tahapan atau tahapan penelitian, seperti dibawah ini:



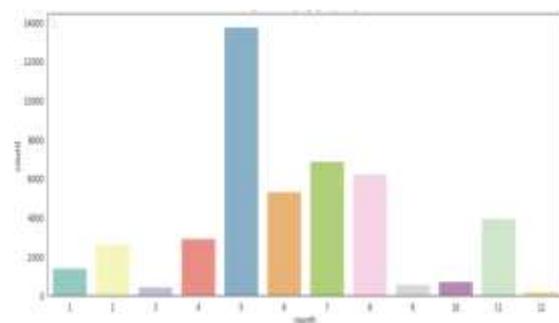
Gambar 4. Tahapan Penelitian

Teknik pengumpulan data yang digunakan oleh penulis adalah dengan pendekatan kualitatif dimana teknik yang digunakan penulis adalah survei literatur akademis pada bidang keilmuan atau kebijakan guna mendapatkan konsep-konsep yang relevan dengan kajian inovasi kebijakan dari masyarakat [13]. Dalam pengumpulan data terdapat 2 (dua) sumber data yaitu sekunder dan primer, data primer diperoleh informasi kunci / informan sedangkan data sekunder diperoleh dari hasil analisis penulis terhadap informasi kunci dan jawaban informan yang telah ditautkan. ke tabel pengkodean dan teori strategis PR. (dokumen dan data jurnal, buku ilmiah dan internet) [14].



Gambar 5. Histogram Dataset

Kemudian jika dilihat tingkatan dari data kampanye dalam satu tahun dapat dilihat pada grafik pada Gambar 3 berikut ini.



Gambar 6. Grafik Jumlah Kampanye per bulan dalam 1 tahun.

Pada tahap ini peneliti melakukan percobaan dan pengujian dengan beberapa metode yaitu Klasifikasi yang meliputi Deep Neural Network (DNN) dengan 2

Hidden Layers, Deep Neural Network (DNN) dengan 3 Hidden Layers, Deep Neural Network (DNN) dengan 4 Hidden Layers, Random Forest, K-nearest neighbour (K-nn). Penulis menambahkan teknik data mining berupa Cross Validation yang bertujuan untuk mendapatkan hasil akurasi yang maksimal. Cara ini sering juga disebut k-fold cross validation, dimana percobaannya dilakukan k kali untuk satu model dengan parameter yang sama.

Dalam percobaan ini peneliti menggunakan investigasi aplikasi Multi Layer Perceptron (MLP) pada Deep Neural Network (DNN), dengan 3 Hidden Layers dengan variasi sebagai berikut:

Tabel 2. Eksperimen Multi Layer Perceptron (MLP) pada DNN dengan 3 Hidden layers

Arsitecture	Activation	Optimizer	Layer (type)	Output Shape	Param	Total Params	Trainable Params	Non-trainable params
Decoder-Encoder	Relu, Sigmoid	Adam	Batch_normalization_60 (batc)	(None, 12)	48	12.199	11.875	324
			Dense_80	(None, 100)	1300			
			batch_normalization_61 (batc)	(None, 100)	400			
			Dropout_60 (dropout)	(None, 100)	0			
			dense_81 (dense)	(None, 50)	5050			
			batch_normalization_62 (batc)	(None, 50)	200			
			Dropout_61 (dropout)	(None, 50)	0			
			dense_82 (dense)	(None, 100)	5100			
			Dropout_62 (dropout)	(None, 100)	0			
			dense_83 (dense)	(None, 1)	101			
Encoder-Decoder			Batch_normalization_60 (batc)	(None, 12)	48	11.499	11.175	324
			Dense_80	(None, 50)	650			
			batch_normalization_61 (batc)	(None, 50)	200			
			Dropout_60 (dropout)	(None, 50)	0			
			dense_81 (dense)	(None, 100)	5100			
			batch_normalization_62 (batc)	(None, 100)	400			
			Dropout_61 (dropout)	(None, 100)	0			
			dense_82 (dense)	(None, 50)	5050			
			Dropout_62 (dropout)	(None, 50)	0			
			dense_83 (dense)	(None, 1)	51			

Hasil pengujian prediksi tingkat kesuksesan promosi bank menggunakan 20;80 data testing dan data training. Dengan menggunakan arsitektur yang baik berdasarkan hasil pengujian, yaitu 3 hidden layer, learning rate dan dilakukan percobaan testing 10 epoch di dapat nilai akurasi terbaik sebesar 90% pada epoch ke-10. Hasil percobaan testing untuk mendapatkan akurasi terbaik ditunjukkan pada table 2.

Tabel 2. Hasil pengujian 10 epoch pada DNN dengan 3 Hidden layers

Epoch	Lr(α)	loss	Binary_acc	vall_loss	acc
1	0,1	0.2537	0.8899	0.2515	0.8882
2	0,1	0.2535	0.8909	0.2521	0.8880
3	0,1	0.2516	0.8901	0.2496	0.8910
4	0,1	0.2502	0.8915	0.2511	0.8884
5	0,1	0.2496	0.8907	0.2542	0.8883
6	0,1	0.2501	0.8910	0.2520	0.8882
7	0,1	0.2499	0.8915	0.2527	0.8901
8	0,1	0.2503	0.8915	0.2522	0.8914
9	0,1	0.2499	0.8923	0.2530	0.8896
10	0,1	0.2475	0.8927	0.2498	0.9005

Tabel 3. Perbandingan Nilai Akurasi Algoritma Klasifikasi

Algoritma	Hasil Akurasi
DNN – 2 Hidden Layers	89%
DNN – 3 Hidden Layers	90%
DNN – 4 Hidden Layers	89%
Random Forest	88%
KNN	89%

IV. SIMPULAN

Dari hasil penelitian yang telah dilakukan algoritma Deep Neural Network (DNN) dengan 3 Hidden Layers memiliki nilai akurasi yang lebih tinggi dibandingkan

dengan menggunakan algoritma K-Nearest Network (KNN) dan algoritma Random Forest. Nilai akurasi untuk model algoritma Deep Neural Network (DNN) dengan 3 Hidden Layers sebesar 90% sedangkan nilai akurasi untuk Deep Neural Network (DNN) dengan 2 Hidden Layers sebesar 89%, Deep Neural Network (DNN) dengan 4 Hidden Layers sebesar 89%, algoritma K-Nearest Network (KNN) sebesar 89% dan algoritma Random Forest sebesar 88%. Sehingga dapat disimpulkan bahwa penerapan algoritma Deep Neural Network (DNN) dengan 3 Hidden Layers mampu menghasilkan tingkat akurasi dalam memprediksi tingkat kesuksesan promosi bank yang lebih baik dibandingkan menggunakan algoritma yaitu algoritma K-Nearest Network (KNN) dan algoritma Random Forest.

DAFTAR PUSTAKA

- [1] E. P. Saputra, "Prediksi Keberhasilan Telemarketing Bank Untuk," *J. Ilmu Pengetah. Dan Teknol. Komput.*, vol. 2, no. 2, pp. 66–72, 2017.
- [2] G. F. Yee, S. F. Sufahani, M. Mamat, M. A. Mohamed, and P. L. Ghazali, "Analog the performance between three classifiers on bank marketing data," *Int. J. Recent Technol. Eng.*,

- vol. 8, no. 2 Special Issue 3, pp. 382–386, 2019, doi: 10.35940/ijrte.B1066.0782S319.
- [3] O. Apampa, “Evaluation of Classification and Ensemble Algorithms for Bank Customer Marketing Response Prediction.,” *J. Int. Technol. Inf. Manag.*, vol. 25, no. 4, pp. 85–100, 2016.
- [4] R. Sulaehani, “Prediksi Keputusan Klien Telemarketing Untuk Deposito Pada Bank Menggunakan Algoritma Naive Bayes Berbasis Backward Elimination,” *Ilk. J. Ilm.*, vol. 8, no. 3, pp. 182–189, 2016, doi: 10.33096/ilkom.v8i3.83.182-189.
- [5] P. A. Wicaksana, I. M. Sudarma, and D. C. Khrisne, “Pengenalan Pola Motif Kain Tenun Gringsing Menggunakan Metode Convolutional Neural Network Dengan Model Arsitektur,” *J. SPEKTRUM*, vol. 6, no. 3, pp. 159–168, 2019.
- [6] L. M. R. Rere and S. Sudjiran, “Optimasi Deep Belief Network Menggunakan Simulated Annealing,” vol. 2, 2018.
- [7] A. Anggoro, A. B. Osmond, and R. E. Saputra, “Recurrent Neural Network Untuk Pengenalan Ucapan Pada Recurrent Neural Network for Speech Recognition on,” vol. 5, no. 3, pp. 6431–6435, 2018.
- [8] T. L. Fine, “Feedforward Neural Network Methodology,” *IEEE Trans. Neural Networks*, vol. 12, no. 3, pp. 647–648, 2001, doi: 10.1109/TNN.2001.925573.
- [9] I. Oktanisa and A. A. Supianto, “Perbandingan Teknik Klasifikasi Dalam Data Mining Untuk Bank a Comparison of Classification Techniques in Data Mining for,” *Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 5, pp. 567–576, 2018, doi: 10.25126/jtiik20185958.
- [10] R. Leonardo, J. Pratama, and C. Chrisnatalis, “Perbandingan Metode Random Forest Dan Naïve Bayes Dalam Prediksi Keberhasilan Klien Telemarketing,” *J. Teknol. Dan Ilmu Komput. Prima*, vol. 3, no. 2, pp. 1–5, 2020.
- [11] S. Dewi, “Pada Prediksi Keberhasilan Pemasaran Produk Layanan Perbankan,” *Techno Nusa Mandiri*, vol. XIII, no. 1, pp. 60–66, 2016.
- [12] W. Darmawan, “Komparasi Metode Data Mining Dalam Memprediksi Nasabah Bank yang akan Memilih Tabungan Deposito Menggunakan Algoritma Klasifikasi,” *IC-Tech*, vol. XIII, no. 1, pp. 49–55, 2018.
- [13] A. Sururi, “Inovasi Kebijakan dalam

- Perspektif Administrasi Publik Menuju Terwujudnya Good Public Policy Governance,” vol. 12, pp. 14–31, 2017, doi: 10.31227/osf.io/6djph.
- [14] R. Rosliana and R. Loisa, “Strategi Cyber Public Relations dalam Memanfaatkan Media Sosial untuk Membangun Citra Perusahaan,” *Prologia*, vol. 2, no. 2, p. 480, 2019, doi: 10.24912/pr.v2i2.3733.
- [15] Wibowo, H., & Indriyani, F. (2018, October). K-Nearest Neighbor Method For Monitoring Of Production And Preservation Information (Treatment) Of Rubber Tree Plant. In *International Conference on Information Technology and Business (ICITB)* (pp. 29-44).

PREDIKSI TINGKAT PELANGGAN *CHURN* PADA PERUSAHAAN TELEKOMUNIKASI DENGAN ALGORITMA ADABOOST

Iqbal Muhammad Latief¹, Agus Subekti², Windu Gata³

¹²³Prodi Ilmu Komputer, Sekolah Tinggi Ilmu Manajemen dan Ilmu Komputer Nusa Mandiri
Jl. Jatiwaringin No. 2, Cipinang Melayu, Makasar Jakarta Timur
Telp. (021) 8005722

e-mail : ¹14002323@nusamandiri.ac.id, ²agus@nusamandiri.ac.id, ³windu@nusamandiri.ac.id

ABSTRACT

With the rapid advancement of the telecommunications industry, and competition between telecommunications companies is increasing, companies need to predict their customers to determine the level of customer loyalty. One of them is by analyzing customer data by doing a Customer Churn Prediction. Predicting Customer Churn is an important business strategy for the company. To acquire new customers is much higher cost than retaining existing customers. The ease of operator switching is one of the serious challenges that the telecommunications industry must face. By predicting customer churn, companies can take immediate action to retain customers. To retain existing customers, the company must improve customer service, improve product quality, and must know in advance which customers have the possibility to leave the company. Prediction can be done by analyzing customer data using data mining techniques. In line with this, gathering information from the telecommunications business can help predict whether customer relationships will leave the company. The data used in this study are secondary data and amount to 7.403 data customers. The data has 21 variables. This study proposes to use the ensemble method namely adaboost, xgboost and random forest and compare them. Algorithm is validated through training data and testing data with a ratio of 80:20. From the results we got using python tools, it was found that the adaboost algorithm has an accuracy of 80%.

Keywords—accuracy, adaboost, churn prediction, compare model, data mining.

ABSTRAK

Dengan pesatnya kemajuan industri telekomunikasi, dan persaingan antar perusahaan telekomunikasi semakin meningkat, perusahaan perlu untuk memprediksi pelanggannya untuk mengetahui tingkat loyalitas pelanggan. Salah satunya adalah dengan menganalisis data pelanggan dengan melakukan prediksi *churn* pelanggan. Prediksi *churn* pelanggan adalah strategi bisnis yang penting bagi perusahaan. Untuk mendapatkan pelanggan baru membutuhkan biaya yang jauh lebih tinggi daripada mempertahankan pelanggan yang sudah ada. Kemudahan operator *switching* merupakan salah satu tantangan serius yang harus dihadapi oleh industri telekomunikasi. Dengan memprediksi pelanggan *churn*, perusahaan dapat secara efektif mengambil keputusan untuk mempertahankannya. Untuk mempertahankan pelanggan yang ada, perusahaan harus meningkatkan layanan pelanggan, meningkatkan kualitas produk, dan harus mengetahui sebelumnya pelanggan mana yang memiliki kemungkinan untuk keluar dari perusahaan. Prediksi dapat dilakukan dengan menganalisis data pelanggan menggunakan teknik *data mining*. Sejalan dengan ini, dengan

mengumpulkan informasi dari bisnis telekomunikasi dapat membantu memprediksi hubungan pelanggan apakah mereka akan meninggalkan perusahaan. Data yang digunakan dalam penelitian ini adalah data sekunder dan berjumlah 7.403 data pelanggan. Data memiliki 21 variabel. Studi ini mengusulkan untuk menggunakan metode *ensemble* yaitu *adaboost*, *xgboost* dan *random forest* dan membandingkannya. Algoritma divalidasi melalui data *training* dan data *testing* dengan rasio 80:20. Dari hasil yang kami dapatkan dengan menggunakan alat bantu python maka ditemukan algoritma *adaboost* memiliki akurasi tertinggi yaitu 80%.

Kata Kunci—*adaboost*, akurasi, data mining, komparasi model, prediksi churn.

I. PENDAHULUAN

Bidang telekomunikasi telah menjadi salah satu bisnis fundamental di negara-negara maju. Seiring dengan itu perusahaan baru mulai bermunculan. Munculnya perusahaan baru ini mengakibatkan persaingan yang semakin ketat. Perusahaan mana pun dapat berkembang jika memiliki jumlah pelanggan yang memadai. Memiliki jumlah pelanggan yang besar memungkinkan pendapatan yang optimal untuk kas perusahaan. Berbagai cara akan dilakukan perusahaan untuk menarik pelanggan tersebut agar menggunakan jasanya. Salah satu cara untuk menarik pelanggan adalah dengan menganalisis data pelanggan yang biasanya disimpan di sejumlah besar database perusahaan. Data ini dapat digunakan untuk menganalisis pelanggan mana yang loyal dan *churn*.

Penelitian tentang *churn* pelanggan semakin penting dan menarik banyak perhatian peneliti [1]. Pelanggan dapat dengan mudah menggunakan haknya

untuk berpindah penyedia layanan dari satu operator ke operator lainnya. Banyaknya operator seluler mendorong persaingan bisnis yang semakin ketat. Kemudahan operator *switching* merupakan salah satu tantangan serius yang harus dihadapi oleh industri telekomunikasi [2]. Mengingat fakta bahwa industri telekomunikasi mengalami tingkat *churn* tahunan rata-rata 30-35 persen, dan biaya untuk merekrut pelanggan baru 5-10 kali lebih mahal daripada mempertahankan pelanggan yang sudah ada, mempertahankan pelanggan menjadi lebih penting daripada mendapatkan pelanggan [3]. *Churn* pelanggan mengacu pada kehilangan pelanggan berkala dalam sebuah organisasi [4].

Industri telekomunikasi berusaha mengembangkan cara untuk memprediksi pelanggan yang berpotensi *churn* sehingga dapat dilakukan tindakan pencegahan karena pengaruh *direct churn* terhadap penurunan pendapatan perusahaan [5]. Aktivitas *churn* sangat berpengaruh terhadap total profit dan citra bisnis,

sehingga sebaiknya dapat diprediksi dan dicegah [6]. Prediksi *churn* dapat digunakan untuk mengidentifikasi *churnes* lebih awal sebelum mereka pindah, dan dapat membantu departemen CRM (*Customer Relationship Management*) untuk mempertahankannya, sehingga potensi kerugian perusahaan dapat dihindari [7]. Dengan demikian penyedia layanan harus mengoptimalkan kinerja model prediksi *churn* dan menerapkan teknik prediksi *churn* serta menerapkan strategi pemasaran yang tepat untuk mempertahankan pelanggan yang ada [8].

Masalah yang dihadapi adalah bagaimana menganalisis pelanggan mana yang loyal atau *churn*. Dalam studi yang dilakukan oleh Keramati disebutkan bahwa untuk bertahan dalam bisnis telekomunikasi harus mampu membedakan antara pelanggan yang memiliki kemungkinan pindah ke pesaing, dan pelanggan yang enggan pindah [9]. Oleh karena itu, prediksi *churn* pelanggan menjadi isu penting dalam bisnis telekomunikasi. Dalam bisnis yang kompetitif, prediktor pelanggan yang andal dianggap sangat berharga.

Proses menghasilkan informasi dari kumpulan data disebut *data mining*. Contoh penelitian yang telah dilakukan dengan *churn analysis* adalah pada perusahaan televisi berlangganan [10],

perusahaan retail [11], dan lain-lain. Diantara metode *data mining* yang telah digunakan untuk menganalisis pelanggan adalah random forest dan decision tree [12] tetapi belum mencapai nilai yang sangat baik. Pada penelitian kali ini akan mencoba untuk mendapatkan akurasi yang lebih baik dari penelitian yang dilakukan sebelumnya yang berjudul *Telecom Customer Churn Prediction* dengan hasil akurasi tertinggi dengan algoritma random forest [12].

Dalam penelitian ini menerapkan teknik klasifikasi *data mining* yang meliputi random forest, adaboost dan xgboost, kemudian membandingkan performanya.

Pada penelitian ini akan mencoba menjawab bagaimana pengklasifikasian data yang paling berpengaruh terhadap tingkat *churn* pada perusahaan telekomunikasi menggunakan metode *ensemble* yaitu random forest, adaboost dan xgboost serta bagaimana akurasi dari ketiga algoritma tersebut.

II. METODE PENELITIAN

Dalam penelitian ini prosedur kerja penelitian ini adalah sebagai berikut:

A) Pemilihan Data

Dataset dalam penelitian ini adalah data sekunder. Dataset tersebut berisi data pelanggan di sebuah perusahaan

telekomunikasi. Dataset ini diunduh dari Kaggle dengan alamat website yaitu <https://www.kaggle.com/blastchar/telco-customer-churn>. Dataset yang digunakan berukuran 955 KB. Dataset ini memiliki 7.403 data pelanggan dan 21 kolom. 21 kolom tersebut adalah variabel yang akan digunakan sebagai prediksi *churn*. Variabel yang terdapat dalam dataset *churn* adalah sebagai berikut [13]:

- A. *Customer ID*: Indeks pelanggan
- B. *Gender*: Jenis kelamin pelanggan. Kolom ini memiliki 2 nilai: *male* dan *female*.
- C. *SeniorCitizen*: Pelanggan adalah warga negara senior. Kolom ini memiliki 2 nilai yaitu 0 dan 1.
- D. *Partners*: Pelanggan memiliki mitra. Kolom ini memiliki 2 nilai, yaitu: *Yes* dan *No*.
- E. *Depedents*: Pelanggan memiliki tanggungan. Kolom ini memiliki 2 nilai, yaitu: *Yes* dan *No*.
- F. *Tenure*: Jumlah bulan pelanggan menggunakan layanan perusahaan.
- G. *PhoneService*: Pelanggan memiliki layanan telepon. Kolom ini memiliki 2 nilai, yaitu *Yes* dan *No*.
- H. *MultipleLines*: Pelanggan memiliki layanan multi saluran. Kolom ini memiliki 2 nilai yaitu *Yes* dan *No*.
- I. *InternetService*: Penyedia layanan internet pelanggan. Kolom ini mempunyai 3 nilai yaitu DSL, *Fiber Optic* dan *No*.
- J. *OnlineSecurity*: Pelanggan memiliki keamanan online. Kolom ini memiliki 3 nilai yaitu *Yes*, *No* dan *No Internet Service*.
- K. *OnlineBackup*: Pelanggan memiliki layanan tampilan *online*. Kolom ini memiliki 3 nilai yaitu *Yes*, *No* dan *No Internet Service*.
- L. *Device Protection*: Pelanggan memiliki layanan perlindungan perangkat. Kolom ini memiliki 3 nilai yaitu *Yes*, *No* dan *No Internet Service*.
- M. *TechSupport*: Pelanggan memiliki dukungan teknis. Kolom ini memiliki 3 nilai yaitu *Yes*, *No* dan *No Internet Service*.
- N. *StreamingTV*: Pelanggan memiliki layanan streaming televisi. Kolom ini memiliki 3 nilai yaitu *Yes*, *No* dan *No Internet Service*.
- O. *StreamingMovies*: Pelanggan memiliki layanan *streaming* film. Kolom ini memiliki 3 nilai yaitu *Yes*, *No* dan *No Internet Service*.
- P. *Contract*: Persyaratan kontrak pelanggan. Kolom ini memiliki 3 nilai yaitu: *Month-to-month*, *One year*, *Two year*.
- Q. *PaperlessBilling*: Pelanggan memiliki tagihan tanpa kertas. Kolom ini memiliki 2 nilai yaitu: *Yes* dan *No*.

R. *PaymentMethod*: Metode pembayaran pelanggan. Kolom ini memiliki 4 nilai yaitu: *Electronic check*, *Mailed check*, *Bank transfer (automatic)* dan *Credit card (automatic)*.

S. *MonthlyCharges*: Ini adalah jumlah yang dibebankan kepada pelanggan setiap bulan.

T. *TotalCharges*: Jumlah total layanan yang dibebankan kepada pelanggan.

U. *Churn*: Kategori pelanggan *churn* atau tidak. Kolom ini memiliki 2 nilai yaitu *Yes* dan *No*.

B) Preprocessing

Setelah data terkumpul, proses selanjutnya adalah proses *preprocessing*. *Preprocessing* data merupakan proses yang bertujuan untuk mengubah data menjadi format yang lebih mudah dan efektif bagi pengguna. Beberapa metode *preprocessing* data yang kami gunakan adalah [14]:

A. Pembersihan data, adalah proses pembersihan data yang memiliki *missing value*.

B. Penyesuaian data, adalah proses menyesuaikan jumlah data untuk setiap target.

C. Pemisahan data, merupakan proses pemisahan data menjadi dua kelompok yaitu *train* dan *test*.

C) Transformation

Transformation adalah proses mengubah data yang dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses ini merupakan proses kreatif dan sangat bergantung pada jenis atau pola informasi yang akan dicari dalam database. Dalam proses ini, data akan dikelompokkan menggunakan metode *One Hot Encoding*.

D) Data mining

Data mining adalah proses mencari pola atau informasi yang menarik dalam data yang dipilih menggunakan teknik atau metode tertentu [15]. Teknik, metode, atau algoritma di dalam *data mining* itu banyak sekali variasinya. Penelitian ini menggunakan algoritma random forest, adaboost dan xgboost.

Algoritma Adaboost

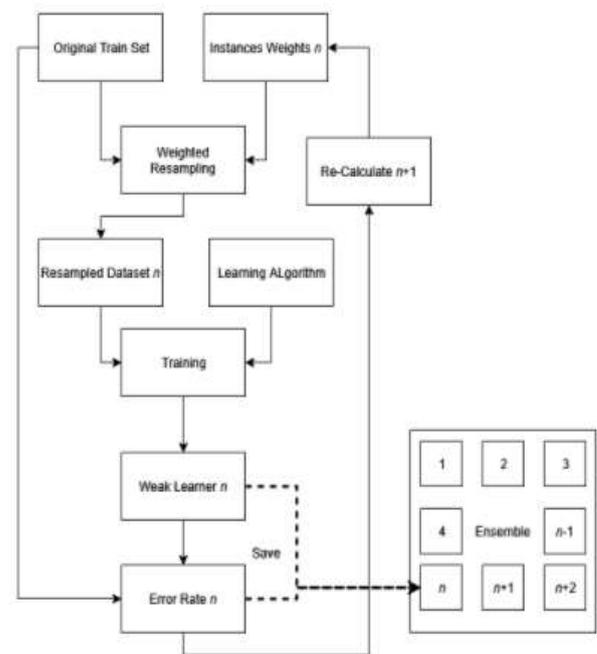
Adaboost adalah salah satu algoritma *boosting* yang paling populer. Mirip dengan *bagging*, ide utama dibalik algoritma adalah membuat sejumlah *weak learners* yang tidak berkolerasi dan kemudian menggabungkan prediksinya. Perbedaan utama dengan *bagging* adalah bahwa alih-alih membuat sejumlah rangkaian-rangkaian *bootstrap* independen, algoritma tersebut secara berurutan melatih setiap *weak learner*, menetapkan bobot ke semua *instance*, dan mengulangi

seluruh proses. Sebagai algoritma *base learner*, biasanya *decision tree* yang terdiri dari satu *node* digunakan. *Decision tree* ini, dengan kedalaman satu tingkat, disebut *decision stumps*.

Algoritma *adaboost* dapat dideskripsikan secara *high-level* dari langkah dasarnya, langkah-langkahnya:

- 1) Inisialisasi semua bobot *instance set train* secara merata, sehingga jumlahnya sama dengan 1.
- 2) Hasilkan set baru dengan pengambilan sampel dengan penggantian, sesuai dengan bobotnya.
- 3) *Train* sebuah *weak learner* pada set sampel.
- 4) Hitung *error* pada *train set* asli.
- 5) Tambahkan *weak learner* ke *ensemble* dan simpan tingkat kesalahannya.
- 6) Sesuaikan bobot, tambah bobot *instance* yang salah diklasifikasikan, dan kurangi bobot *instance* yang diklasifikasikan dengan benar.
- 7) Ulangi dari langkah 2.
- 8) *Weak learners* digabungkan dengan *voting*. *Vote* setiap *learner* diberi bobot sesuai tingkat kesalahannya [16].

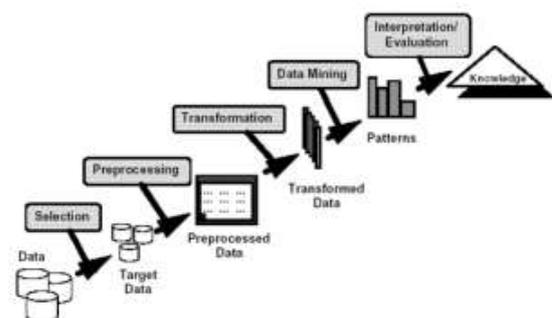
Seluruh proses digambarkan dalam gambar berikut:



Gambar 1. Langkah-Langkah Algoritma Adaboost

E) Interpretasi/Evaluasi

Tahapan mulai dari menyeleksi data yaitu melihat *missing value*, kemudian melakukan transformasi ke bentuk *numeric* kemudian mencari mendapatkan model terbaik dan terakhir pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dipahami oleh pihak yang berkepentingan. Pada tahap ini, pengetahuan yang dihasilkan dari *data mining* akan dirilis.



Gambar 2. Proses Data Mining

III. HASIL

Objek yang digunakan dalam penelitian ini adalah data pelanggan pada perusahaan telekomunikasi. Data pelanggan diperoleh dengan mengunduh di Kaggle pada alamat <https://www.kaggle.com/blastchar/telco-customer-churn>. Dataset ini berisi 7.043 data pelanggan. Berikut 5 contoh data pelanggan dalam dataset tersebut.

	customerID	gender	SeniorCitizen	Partner	Dependents
0	7590-VHVEG	Female	0	Yes	No
1	5575-GNVDE	Male	0	No	No
2	3668-QPYBK	Male	0	No	No
3	7795-CFOCW	Male	0	No	No
4	9237-HQITU	Female	0	No	No

Gambar 3. Contoh 5 Data Pelanggan dari Dataset

tenure	PhoneService	MultipleLines	InternetService	Online Security
1	No	No phone service	DSL	No
34	Yes	No	DSL	Yes
2	Yes	No	DSL	Yes
45	No	No phone service	DSL	Yes
2	Yes	No	Fiber optic	No

Gambar 4. Lanjutan Contoh 5 Data Pelanggan dari Dataset

DeviceProtection	TechSupport	StreamingTV	StreamingMovies	Contract
No	No	No	No	Month-to-month
Yes	No	No	No	One year
No	No	No	No	Month-to-month
Yes	Yes	No	No	One year
No	No	No	No	Month-to-month

Gambar 5. Lanjutan Contoh 5 Data Pelanggan dari Dataset

PaperlessBilling	PaymentMethod	MonthlyCharges	TotalCharges	Churn
Yes	Electronic check	29.85	29.85	No
No	Mailed check	56.95	1889.5	No
Yes	Mailed check	53.85	108.15	Yes
No	Bank transfer (automatic)	42.30	1840.75	No
Yes	Electronic check	70.70	151.65	Yes

Gambar 6. Lanjutan Contoh 5 Data Pelanggan dari Dataset

Pertama menemukan *missing values* dari data, kemudian mengisinya dengan 0 (nol). Terdapat 11 data *missing values* pada variabel *TotalCharges*.

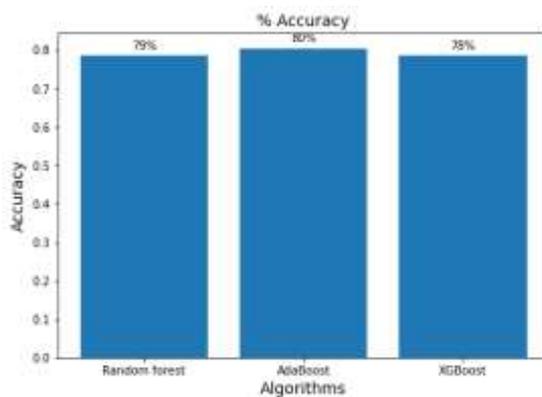
Sasaran pada penelitian ini adalah kolom *Churn*. Jumlah data untuk target “No” adalah 73,4% dan “Yes” adalah 26,6%.

Sebelum membentuk aturan klasifikasi, perlu dilakukan pembagian data menjadi 2 kelompok yaitu *train* dan *test*. Berbagi data ini bertujuan untuk menganalisis apakah aturan klasifikasi yang dihasilkan oleh algoritma adaboost dapat digunakan untuk memprediksi

churn. Distribusi dataset ini menggunakan rasio 80% data *training* dan 20% data *testing*.

Kemudian setelah itu dilakukan pemodelan dengan algoritma random forest, adaboost dan xgboost dengan parameter *default*.

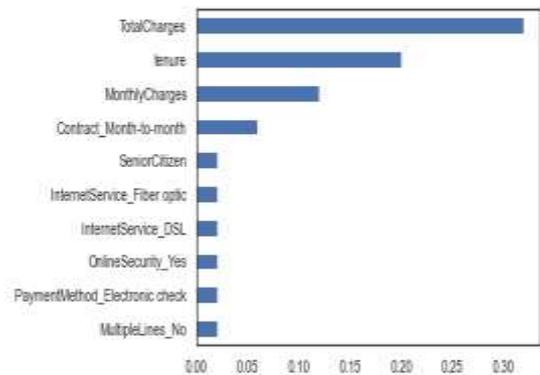
Setelah pemodelan kemudian dilakukan penghitungan akurasi dari masing-masing algoritma, maka didapat hasil seperti gambar 7.



Gambar 7. Perbandingan Ketiga Algoritma

Dari gambar 7 dapat dilihat bahwa adaboost memiliki akurasi tertinggi yaitu 80%.

Dari Algoritma adaboost, *Total Charges*, *tenure*, *MonthlyCharges*, adalah variabel prediktor paling penting untuk memprediksi *Churn*.



Gambar 8. Variabel Penting yang Mempengaruhi *Churn* dengan Algoritma Adaboost

Akurasi yang dihasilkan dari pengujian aturan klasifikasi ini dari data pengujiannya adalah 80%.

Dari aturan klasifikasi yang dibentuk menggunakan algoritma adaboost, variabel *TotalCharges* memiliki pengaruh cukup kuat terhadap prediksi *churn* ini.

Dengan mengetahui bahwa variabel *TotalCharges* memiliki pengaruh yang cukup besar dalam memprediksi *churn*, maka perusahaan dapat melakukan promosi yang lebih baik kepada pelanggan dengan adanya *TotalCharges* yang berpotensi *churn*.

IV. KESIMPULAN

Setelah melakukan penelitian ini, dapat ditarik kesimpulan bahwa algoritma adaboost dapat memprediksi masalah *churn* lebih baik dari algoritma random forest dan xgboost serta *TotalCharges* adalah fitur yang paling penting dalam

memprediksi *churn* dengan tingkat akurasi 80% dari pada penelitian sebelumnya dengan algoritma random forest [12].

DAFTAR PUSTAKA

- [1] E. B. Lee, J. Kim, and S. G. Lee, "Predicting customer churn in mobile industry using data mining technology," *Ind. Manag. Data Syst.*, vol. 117, pp. 90–109, 2017, doi: 10.1108/IMDS-12-2015-0509.
- [2] B. Huang, M. T. Kechadi, and B. Buckley, "Customer churn prediction in telecommunications," *Expert Syst. Appl.*, vol. 39, 2012, doi: 10.1016/j.eswa.2011.08.024.
- [3] J. Lu and D. Ph, "Predicting Customer Churn in the Telecommunications Industry — An Application of Survival Analysis Modeling Using SAS," *Techniques*, 2002.
- [4] A. Churi, M. Divekar, S. Dashpute, and P. Kamble, "Analysis of customer churn in mobile industry using data mining.," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 5(3), pp. 225–230, 2015.
- [5] A. K. Ahmad, A. Jafar, and K. Aljoumaa, "Customer churn prediction in telecom using machine learning in big data platform," *J. Big Data*, 2019, doi: 10.1186/s40537-019-0191-6.
- [6] J. Pamina, T. Dhiliphan Rajkumar, S. Kiruthika, T. Suganya, and F. Femila, "Exploring hybrid and ensemble models for customer churn prediction in telecom sector," *Int. J. Recent Technol. Eng.*, 2019, doi: 10.35940/ijrte.A9170.078219.
- [7] V. Umayaparvathi and K. Iyakutti, "Applications of Data Mining Techniques in Telecom Churn Prediction," *Int. J. Comput. Appl.*, 2012, doi: 10.5120/5814-8122.
- [8] R. Misra, S. Singh, and R. Mahajan, "An empirical study on the cellular subscribers churn, selection factors and satisfaction with the services," 2019, doi: 10.1504/ijpd.2019.10020377.
- [9] A. Keramati, R. Jafari-Marandi, M. Aliannejadi, I. Ahmadian, M. Mozaffari, and U. Abbasi, "Improved churn prediction in telecommunication industry using data mining techniques," *Appl. Soft Comput. J.*, 2014, doi: 10.1016/j.asoc.2014.08.041.
- [10] N. Suryana, "PREDIKSI CHURN DAN SEGMENTASI PELANGGAN TV BERLANGGANAN (STUDI KASUS TRANSVISION JAWA BARAT)," *J. TEDC*, vol. 11(2), no.

- Vol 11 No 2 (2017): Jurnal TEDC, pp. 185–191, 2019, [Online]. Available: <http://ejournal.poltektedc.ac.id/index.php/tedc/article/view/77>.
- [11] N. W. Wardani, G. R. Dantes, and G. Indrawan, “Prediksi Customer Churn dengan Algoritma Decision Tree C4.5 Berdasarkan Segmentasi Pelanggan untuk Mempertahankan Pelanggan pada Perusahaan Retail,” *J. Resist. (Rekayasa Sist. Komputer)*, 2018, doi: 10.31598/jurnalresistor.v1i1.219.
- [12] M. Manasa, M. N. Reddy, M. L. Sahithi, P. Y. Kumar, and V. Sandhya, “Telecom Customer Churn Prediction,” vol. 8, 2020, doi: 10.22214/ijraset.2020.5479.
- [13] Kaggle, “Telco Customer Churn: Focused customer retention programs,” 2018. <https://www.kaggle.com/blastchar/telco-customer-churn> (accessed Nov. 06, 2020).
- [14] R. R. Rerung, “Penerapan Data Mining dengan Memanfaatkan Metode Association Rule untuk Promosi Produk,” *J. Teknol. Rekayasa*, 2018, doi: 10.31544/jtera.v3.i1.2018.89-98.
- [15] E. Srikanti, R. F. Yansi, Norvahina, I. Permana, and F. N. Salisah, “Penerapan Data Mining Untuk Menganalisis Penjualan Barang dengan Menggunakan Metode Apriori pada Supermarket Sejahtera Lhoksumawe,” *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, 2018.
- [16] G. Kyriakides, *Hands-On Ensemble Learning with Python: Build Highly Optimized Ensemble Machine Learning Models Using Scikit-Learn and Keras*. Birmingham: Packt Publishing Ltd., 2019.
- [17] Sabita, H., & Herwanto, R. (2020). Pantauan Prediktif Covid-19 Dengan Menggunakan Metode SIR dan Model Statistik Di Indonesia. *TEKNIKA*, 14(2), 145-150.

PENERAPAN DATA MINING TERHADAP DATA COVID-19 MENGUNAKAN ALGORITMA KLASIFIKASI

Rizka Dahlia¹, Nanik Wuryani², Sri Hadiani³, Windu Gata⁴, Arina Selawati⁵

¹²³⁴Fakultas Ilmu Komputer, Sekolah Tinggi Ilmu Komputer Nusa Mandiri Jakarta
5, Jl. Kramat Raya No. 18 RT. 05 RW 7, Kwitang, Kec. Senen
Kota Jakarta Pusat DKI Jakarta 10450

⁵Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika
2, Jl. Kramat Raya No.98, RT.2/RW.9, Kwitang, Kec. Senen
Kota Jakarta Pusat, DKI Jakarta 10450

e-mail : 14002335@nusamandiri.ac.id¹, 14002336@nusamandiri.ac.id², 1sri.shv@nusamandiri.ac.id³,
windu@nusamandiri.ac.id⁴, arina.asq@bsi.ac.id

ABSTRACT

Coronavirus 2019 or more commonly referred to as COVID-19 is a type of virus that attacks the respiratory system. Until now the number of spread and the number of deaths caused by this virus continues to increase. As of April 21, 2020, based on data from the WHO, the total number of cases infected with this virus reached 2,397,217 with 162 deaths from all over the world. For South Korea itself, as of March 21, 2020, the total number of infected cases was 10,683 with a total of 237 deaths. In this study, researchers conducted data processing on the spread of COVID-19 in South Korea with Rapidminer using a classification algorithm, namely Naïve Bayes, C4.5, and K-Nearest Neighbor by performing the stages of selection, preprocessing, transformation, data mining and interpretation or evaluating the quality of the best accuracy of 80.79% with AUC of 0.881 achieved by the Naïve Bayes algorithm. The distribution of the data found that the influential attribute of the isolated class factor from the patient contained in the sex attribute where more women experienced isolation.

Keywords— COVID-19, data mining, classification, C4.5, Naïve Bayes, K-NN

ABSTRAK

Coronavirus 2019 atau yang lebih sering disebut dengan COVID-19 merupakan salah satu virus jenis yang menyerang sistem pernapasan. Hingga saat ini angka penyebaran dan angka kematian yang diakibatkan virus ini terus bertambah. Per tanggal 21 April 2020, berdasarkan data dari WHO total kasus yang terinfeksi virus ini mencapai 2.397.217 dengan kasus kematian mencapai 162 kasus dari seluruh dunia. Untuk Korea Selatan sendiri, per tanggal 21 Maret 2020 total kasus yang terinfeksi mencapai 10.683 dengan total kematian sebanyak 237. Pada penelitian kali ini peneliti melakukan pengolahan data penyebaran COVID-19 di Korea Selatan dengan Rapidminer menggunakan algoritma klasifikasi yaitu Naïve Bayes, C4.5, dan K-Nearest Neighbor dengan melakukan tahapan selection, preprocessing, transformation, data mining dan interpretation atau evaluation menghasilkan akurasi terbaik 80,79% dengan AUC 0.881 yang diraih oleh algoritma Naïve Bayes. Distribusi data yang didapatkan bahwa atribut yang berpengaruh dari faktor class isolated dari pasien terdapat pada atribut sex dimana female lebih banyak yang mengalami isolated.

Kata Kunci— COVID-19, data mining, klasifikasi, C4.5, Naïve Bayes, K-NN

I. PENDAHULUAN

Penyakit *Coronavirus 2019* atau lebih dikenal dengan istilah *COVID-19* merupakan suatu wabah yang awalnya terdeteksi di Kota Wuhan, Cina pada Desember 2019. Sebelum disebut sebagai *COVID-19*, WHO atau *World Health Organization* memberikan nama sementara virus baru ini sebagai *Coronavirus Novel 2019 (2019-nCoV)*. Dan pada 21 April 2020 WHO secara resmi menyebut virus 2019-nCoV menjadi *COVID-19* [1][2].

COVID-19 bermula dari *betacoronavirus (SARS-CoV-2)* yang menyerang bagian saluran pernapasan bagian bawah yang berubah menjadi pneumonia di tubuh manusia. Virus *COVID-19* merupakan *coronavirus* jenis baru. *COVID-19* dianggap sebagai kerabat dari *Severe Acute Respiratory Syndrome (SARS)* dan *Middle East Respiratory Syndrome Coronavirus (MERS)* [1].

Berdasarkan data yang didapat dari WHO, terdapat 179 negara yang sudah terpapar virus *COVID-19*. Hal menandakan bahwa virus ini memiliki tingkat paparan yang sangat tinggi dan cepat. Cara penyebarannya juga sangat sederhana. Penyebarannya dapat berupa bersin, batuk, atau berinteraksi dengan orang yang sudah terinfeksi. Dan virus ini

lebih rentan terhadap orang tua dan mereka yang memang sudah memiliki riwayat penyakit serius.

Per tanggal 21 April 2020, berdasarkan data dari WHO total kasus yang terinfeksi virus ini mencapai 2.397.217 dengan kasus kematian mencapai 162 kasus dari seluruh dunia. Untuk Korea Selatan sendiri, per tanggal 21 Maret 2020 total kasus yang terinfeksi mencapai 10.683 dengan total kematian sebanyak 237. Data ini masih melakukan pembaruan hingga saat ini dan jumlah yang terinfeksi, sembuh serta meninggal akan bertambah. Ada beberapa faktor yang mempengaruhi cepatnya penyebaran virus ini yaitu umur tua, banyaknya orang bepergian ke negara yang sudah terinfeksi, melakukan kontak dengan orang yang terinfeksi, dan sebagainya [3]. Faktor-faktor tersebut dapat menjadi data dan dapat diolah dengan *data mining*.

Penelitian terdahulu yang menggunakan *dataset Data Science for COVID-19 (DS4C)* yang diambil dari kaggle yang juga digunakan pada penelitian ini pernah dilakukan oleh Al-Najjar dan Al-Rousan membahas mengenai prediksi kesembuhan dan kematian pasien *Covid-19* di Korea Selatan dengan algoritma yang digunakan yaitu *Artificial Neural Network (ANN)* dengan hasil yang didapatkan pada

penelitian ini adalah usulan untuk memperhatikan penyebab infeksi untuk peningkatan kesembuhan pasien Covid-19 dan pengendalian regional atau daerah untuk meminimalisir kematian[4].

Penelitian dengan *dataset* yang sama dilakukan juga oleh Muhammad dkk. yang membahas tentang penggunaan beberapa model untuk mendapatkan akurasi tertinggi. Model yang digunakan antara lain *Decision Tree*, *Support Vector Machine*, *Naïve Bayes*, *Logistic Regression*, *Random Forest*, dan *K-Nearest Neighbor*. Dengan akurasi paling tinggi yang didapatkan dari beberapa model yang digunakan adalah *Decision Tree* memiliki akurasi yang tertinggi dengan akurasi 99.85% [5]. Kekurangan dari penelitian ini adalah terpaku terhadap akurasi tanpa melihat matriks yang mempengaruhi baik atau buruknya sebuah model yang dibuat.

Melihat dari penelitian sebelumnya, penelitian yang akan dilakukan kali ini adalah melakukan klasifikasi dengan atribut yang digunakan yaitu *sex*, *age*, *city*, *infaction case* serta *state* dengan menggunakan model klasifikasi *Naïve Bayes*, *C4.5*, dan *K-Nearest Neighbor* dengan tujuan menghasilkan klasifikasi pasien Covid-19 di Korea Selatan dengan menguji atribut yang berhubungan dengan penentuan status pasien Covid-19.

1. *Naïve Bayes*

Naïve Bayes merupakan salah satu algoritma *data mining* yang terdapat pada klasifikasi yang di ambil dari nama seorang ahli matematika yang bernama *Thomas Bayes* dan merupakan seorang menteri *Prebysterian* Inggris. Algoritma *Naïve Bayes* menggunakan teknik percabangan matematika dengan mencari peluang terbesar dari kemungkinan dalam klasifikasi berdasarkan frekuensi tiap klasifikasi terhadap data *training* yang sering di sebut dengan teori probabilitik [6]. Adapun rumus perhitungan dari *Naïve Bayes* adalah sebagai berikut [7]:

$$P(X|Y) = \frac{P(Y|X) \times P(X)}{P(Y)} \quad (1)$$

Keterangan :

- Y = data dengan kelas yang belum diketahui
- X = hipotesis data Y merupakan suatu kelas spesifik
- P(X|Y) = probabilitas hipotesis X berdasarkan kondisi Y
- P(X) = Probabilitas hipotesis X
- P(Y|X) = Probabilitas Y berdasarkan kondisi pada hipotesis X
- P(Y) = Probabilitas Y

2. C4.5

C4.5 merupakan salah satu algoritma pada metode klasifikasi dengan menghasilkan sebuah pohon keputusan. Pohon keputusan berisikan aturan yang direpresentasikan dengan bahasa yang mudah dimengerti. Dalam membuat

sebuah pohon keputusan dibutuhkan atribut yang dipilih sebagai akar kemudian akan membentuk cabang setiap nilai. Proses itu akan diulang terus menerus hingga setiap cabang memiliki kelas yang sama [8]. Proses tersebut memiliki perhitungan yaitu sebagai berikut:

$$Gain(S, A) = Entropy(s) \sum n * Entropy(s) \quad (2)$$

Keterangan:

S = Himpunan Kasus

A = atribut

n = jumlah partisi dalam atribut

|Si| = jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

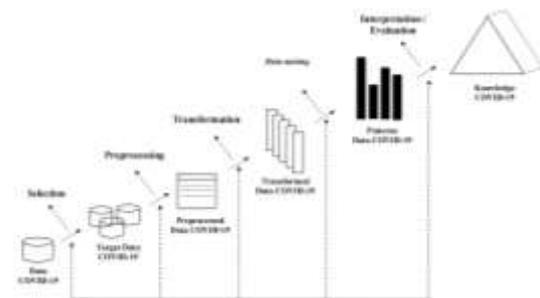
3. k-NN

k-NN atau *K-Nearest Neighbor* merupakan salah satu algoritma yang di pakai dalam melakukan klasifikasi. Algoritma k-NN melakukan klasifikasi dengan cara mencari *class* terdekat dengan jumlah data berupa *k* dengan class yang lain. *Class* terdekat dengan jumlah *k* terbesar dipilih sebagai *class* untuk diprediksi dengan *class* yang baru [9]. Untuk mencari dekat atau jauhnya sebuah *class* dicari dengan persamaan sebagai berikut.

$$d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \quad (3)$$

II. METODE PENELITIAN

Metode atau tahapan penelitian sangat diperlukan ketika melakukan penelitian. Tahap penelitian dilakukan secara sistematis guna membantu penelitian agar terarah dalam melakukan penelitian. Adapun tahapan penelitian yang dilakukan pada penelitian ini sebagai berikut:



Gambar 1. Tahapan Penelitian Data mining COVID-19 Algoritma Klasifikasi

a. Selection

Sebelum melakukan proses *selection* lakukan terlebih dahulu pengumpulan data. Pengumpulan data dilakukan dengan mencari *dataset* pada *repository* atau jika memiliki data sendiri yang terdapat pada perusahaan jika memang dibutuhkan dalam penelitian. Kemudian pilih data yang sesuai dengan pembahasan yang ingin dibahas.

b. Preprocessing

Tahap preprocessing merupakan tahap di mana data yang sudah diseleksi kemudian dihilangkan berbagai macam data yang tidak memiliki isi atau

kandungan. Proses ini dilakukan agar mendapatkan kualitas data yang baik.

c. Transformation

Pada tahap ini data yang telah melalui tahap *preprocessing* akan dilanjutkan pemilihan atribut yang tepat untuk melakukan penelitian. Atribut ini dipilih untuk mempermudah dalam melakukan proses *data mining*.

d. Data mining

Proses *data mining* merupakan proses dimana algoritma yang sudah ditentukan untuk diterapkan pada *RapidMiner* dengan menggunakan data-data yang sudah diolah pada proses sebelumnya serta menerapkan model yang diinginkan. Pada Proses *data mining* menggunakan *RapidMiner* akan menghasilkan sebuah keputusan atau pola berdasarkan algoritma yang ingin digunakan.

e. Interpretation / Evaluation

Pola atau keputusan yang didapat pada proses *data mining* disajikan dalam bentuk yang mudah dimengerti. Hal ini dilakukan untuk mempermudah seseorang dalam membaca dan melakukan koreksi hasil yang telah didapat.

III. HASIL DAN PEMBAHASAN

Dalam penelitian yang menggunakan Data COVID-19 didapatkan di *repository kaggle.com* dengan nama *Science for COVID-19 (DS4C)*. Terdapat beberapa data yang disediakan. Namun *dataset* yang diambil untuk melakukan penelitian ini terfokus kepada *dataset pasien info*. *Dataset* ini berisikan informasi mengenai pasien yang terpapar COVID-19 di Korea Selatan. Berikut adalah bentuk data dari *dataset pasien info*.

Tabel 1. Contoh Data Pasien info

patient_id	global_num	sex	birth_year	age	country	province	city	disease	infection_case	infection_order	infectd_by	contact_number	symptom_onset_date	confirmed_date	released_date	deceased_date	state
1000000001	2	male	1964	50s	Korea	Seoul	Gangseo-gu		overseas inflow	1		75	1/22/2020	1/23/2020	2/5/2020		released
1000000002	5	male	1987	30s	Korea	Seoul	Jungnang-gu		overseas inflow	1		31		1/30/2020	3/2/2020		released
1000000003	6	male	1964	50s	Korea	Seoul	Jongno-gu		contact with patient	2	2002000001	17		1/30/2020	2/19/2020		released
1000000004	7	male	1991	20s	Korea	Seoul	Mapo-gu		overseas inflow	1		9	1/26/2020	1/30/2020	2/15/2020		released
1000000005	9	female	1992	20s	Korea	Seoul	Seongbuk-gu		contact with patient	2	1000000002	2		1/31/2020	2/24/2020		released
1000000006	10	female	1966	50s	Korea	Seoul	Jongno-gu		contact with patient	3	1000000003	43		1/31/2020	2/19/2020		released
1000000007	11	male	1995	20s	Korea	Seoul	Jongno-gu		contact with patient	3	1000000003	0		1/31/2020	2/10/2020		released
1000000008	13	male	1992	20s	Korea	Seoul	etc		overseas inflow	1		0		2/2/2020	2/24/2020		released
1000000009	19	male	1983	30s	Korea	Seoul	Songpa-gu		overseas inflow	2		68		2/5/2020	2/21/2020		released
1000000010	21	female	1960	60s	Korea	Seoul	Seongbuk-gu		contact with patient	3	1000000003	6		2/5/2020	2/29/2020		released
1000000011	23	female	1962	50s	China	Seoul	Seodaemun-gu		overseas inflow	1		23		2/6/2020	2/29/2020		released
1000000012	24	male	1992	20s	Korea	Seoul	etc		overseas inflow	1		0		2/7/2020	2/27/2020		released

Data tersebut memiliki beberapa atribut yang tidak diperlukan dalam proses pengolahan. Maka dari itu proses selanjutnya adalah memilih atribut yang dibutuhkan. Atribut yang digunakan dalam penelitian ini adalah *sex*, *age*, *city*, *infection case*, dan *state*. Atribut *state* berfungsi sebagai label keputusan yang akan dihasilkan. Terdapat dua *class* pada atribut *state* yang digunakan yaitu *released* dan *isolated*.

atau *blank*. Data yang didapatkan sebanyak 2420 data setelah melalui proses *sorting*. Dan data tersebut yang diterapkan di *Rapidminer* untuk proses pembuatan model *data mining* menggunakan algoritma C4.5, *Naïve Bayes*, dan k-NN.

Tabel 2. Contoh Data Pasien info Setelah Pemilihan Atribut

sex	age	city	infection_case	state
male	50s	Gangseo-gu	overseas inflow	released
male	30s	Jungnang-gu	overseas inflow	released
male	50s	Jongno-gu	contact with patient	released
male	20s	Mapo-gu	overseas inflow	released
female	20s	Seongbuk-gu	contact with patient	released
female	50s	Jongno-gu	contact with patient	released
male	20s	Jongno-gu	contact with patient	released
male	20s	etc	overseas inflow	released
male	30s	Songpa-gu	overseas inflow	released
female	60s	Seongbuk-gu	contact with patient	released
female	50s	Seodaemun-gu	overseas inflow	released
male	20s	etc	overseas inflow	released
male	80s	Jongno-gu	contact with patient	released
female	60s	Jongno-gu	contact with patient	released
male	70s	Seongdong-gu	Seongdong-gu APT	isolated
male	70s	Jongno-gu	contact with patient	released
male	70s	Jongno-gu	contact with patient	released
male	20s	etc	etc	isolated
female	70s	Jongno-gu	contact with patient	released



Gambar 2. Model Rapidminer Algoritma Klasifikasi

Setelah data dimasukkan ke dalam *RapidMiner*, jalankan model *RapidMiner*. Setelah *RapidMiner* dijalankan, maka akan tampil hasil sebagai berikut.

	True released	True isolated	State precision
pred released	155	47	77.83%
pred isolated	40	225	83.83%
state total	79.30%	82.72%	

Gambar 3. Hasil Akurasi Naïve Bayes

Dari proses *Rapidminer* menggunakan algoritma *Naïve Bayes* menghasilkan akurasi sebesar 80,79% yang ditunjukkan pada gambar diatas.

Setelah pemilihan atribut, kemudian menghapus data kosong. Proses menghilangkan data kosong ini dilakukan dengan melakukan *sorting* data kosong

	Real released	Real isolated	Class prediction
pred released	0	0	0.00%
pred isolated	252	272	56.20%
class real	0.00%	100.00%	

Gambar 4. Hasil Akurasi C4.5

Untuk algoritma C4.5 mendapatkan hasil akurasi yang lebih kecil dari *Naïve Bayes* dengan hasil 54,20%.

	Real released	Real isolated	Class prediction
pred released	156	156	60.74%
pred isolated	26	170	77.31%
class real	82.98%	62.60%	

Gambar 5. Hasil Akurasi k-NN

Sedangkan hasil yang didapatkan pada algoritma k-NN mendapatkan akurasi 60,74 %. Perbandingan akurasi dari ketiga algoritma dapat dilihat pada tabel berikut.

Tabel 3. Perbandingan Akurasi Algoritma C4.5, *Naïve Bayes*, dan k-NN

Algoritma	Akurasi	AUC
<i>Naïve Bayes</i>	80.79%	0.881
C4.5	56.20%	0.743
k-NN	60.74%	0.567

Dari tabel di atas dapat dilihat untuk nilai akurasi yang dicapai dari ketiga algoritma. *Naïve Bayes* memiliki akurasi 80,79% dengan AUC 0.881, algoritma C4.5 mendapatkan akurasi 56,20% dengan AUC 0.743 dan k-NN mendapatkan akurasi 60,74% dengan AUC 0.567. Jika ketiga algoritma tersebut dibandingkan maka nilai akurasi dan AUC terbaik

adalah algoritma *Naïve Bayes*. Berikut adalah penyebaran data dari algoritma *Naïve Bayes*.

Tabel 4. Penyebaran Data pada Algoritma *Naïve Bayes*

Attribute	Parameter	released	isolated
sex	value=female	0.555	0.560
sex	value=male	0.445	0.440
infection_case	value=contact with patient	0.346	0.388
infection_case	value=overseas inflow	0.074	0.305
age	value=20s	0.251	0.233
infection_case	value=etc	0.279	0.175
age	value=50s	0.177	0.168
age	value=30s	0.163	0.144
age	value=60s	0.096	0.139
age	value=40s	0.187	0.132
city	value=Seongnam-si	0.000	0.086
age	value=70s	0.037	0.069
age	value=10s	0.038	0.048
city	value=Gyeongsan-si	0.033	0.046
age	value=80s	0.026	0.043

IV. SIMPULAN

Hasil dari klasifikasi pada penelitian ini dengan atribut *sex*, *age*, *city*, *infection case*, dan *state* menggunakan algoritma C4.5, *Naïve Bayes*, dan k-NN pada *Rapidminer* menghasilkan banyak pasien masuk ke class *isolated*. Dari ketiga algoritma, didapatkan akurasi tertinggi dari algoritma *Naïve Bayes* dengan hasil 80,79% dengan hasil *isolated* dan memiliki AUC 0.881 yang dikategorikan sebagai *good clasification* atau klasifikasi yang baik. Adapun hasil distribusi data yang didapatkan bahwa atribut yang berpengaruh dari faktor class *isolated* dari

pasien terdapat pada atribut *sex* dimana *female* lebih banyak yang mengalami *isolated* dari algoritma *Naïve Bayes*.

DAFTAR PUSTAKA

- [1] C. Sohrabi *et al.*, “World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19),” *International Journal of Surgery*. 2020, doi: 10.1016/j.ijssu.2020.02.034.
- [2] P. Sun, X. Lu, C. Xu, W. Sun, and B. Pan, “Understanding of COVID-19 based on current evidence,” *Journal of Medical Virology*. 2020, doi: 10.1002/jmv.25722.
- [3] S. A. Rasmussen, J. C. Smulian, J. A. Lednicky, T. S. Wen, and D. J. Jamieson, “Coronavirus Disease 2019 (COVID-19) and Pregnancy: What obstetricians need to know.,” *Am. J. Obstet. Gynecol.*, 2020, doi: 10.1016/j.ajog.2020.02.017.
- [4] H. Al-Najjar and N. Al-Rousan, “A classifier prediction model to predict the status of Coronavirus CoVID-19 patients in South Korea,” *Eur. Rev. Med. Pharmacol. Sci.*, vol. 24, no. 6, pp. 3400–3403, 2020, doi: 10.26355/eurrev_202003_20709.
- [5] L. J. Muhammad, M. M. Islam, S. S. Usman, and S. I. Ayon, “Predictive *Data mining* Models for Novel Coronavirus (COVID-19) Infected Patients’ Recovery,” *SN Comput. Sci.*, vol. 1, no. 4, pp. 1–7, 2020, doi: 10.1007/s42979-020-00216-w.
- [6] R. A. Saputra, “Komparasi Algoritma Klasifikasi *Data mining* Untuk Memprediksi Penyakit Tuberculosis (Tb): Studi Kasus Puskesmas Karawang,” *Semin. Nas. Inov. dan Tren*, 2014.
- [7] W. D. Septiani, “Komparasi Metode Klasifikasi *Data mining* Algoritma C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis,” *J. Pilar Nusa Mandiri*, vol. 13, no. 1, pp. 76–84, 2017.
- [8] F. Parung, “Penerapan algoritma decision tree c4.5 dalam penerimaan guru pada smk sirajul falah parung,” vol. 11, no. 2, pp. 192–198, 2018.
- [9] I. G. Harsemadi, M. Sudarma, and N. Pramaita, “Implementasi Algoritma K-Nearest Neighbor pada Perangkat Lunak Pengelompokan Musik untuk Menentukan Suasana Hati,” *Maj. Ilm. Teknol. Elektro*, vol. 16, no. 1, pp. 14–20, 2017, doi: 10.24843/mite.1601.03.

- [10] Sabita, H., & Herwanto, R. (2020).
Pantauan Prediktif Covid-19 Dengan
Menggunakan Metode SIR dan
Model Statistik Di
Indonesia. *TEKNIKA*, 14(2), 145-
150.

VIDEO RANCANGAN BY-PRO (SISTEM PENGUKURAN DAN PENDATAAN BALITA DI POSYANDU)

Chyntia Raras Ajeng Widiawati¹, Indra Alan Nugroho², Irfan Rifai Aziz³, Rizky Arsyansyah⁴

¹²³⁴Fakultas Ilmu Komputer, Informatika Universitas Amikom Purwokerto
Jl. Letjend Pol. Soemarto, Watumas, Purwanegara, Kec. Purwokerto Tim., Kabupaten Banyumas,
Jawa Tengah 53127
e-mail : ¹chyntiaraw@amikompurwokerto.ac.id, ²indraalan3@gmail.com,
³rifaairfan41@gmail.com, ⁴rarsyansyahr@gmail.com

ABSTRACT

Integrated Healthcare Center (IHC) activity is important thing to do to make sure the toddlers health by measuring toddler's weight and height. Recently IHC activity is done by using manual measurement tools that is sometimes dangerous for the toddler, and it's quite troubled for IHC cadre in reporting and organizing data. It's needed in designing product that can support continuity of IHC activities so it can run safety and comfortably. This research will design a product BY-PRO, a main solution to solve the problems. The design consist of Arduino microcontroller as data processor, cell load sensor (weight), optocoupler modul (height), and sensor K02 (touch sencor) connected to Android via modul (Bluetooth) HC-05. Android Application can be used by IHC cadre to show identity data and toddler measurement saved to database and distributed to web-based Information System to monitor realtime health by midwife. The design result is presented in a video, so it can give an overview of the designed product.

Keywords— *Toddler Height Measuring, Toddler Weight Measuring, Android, Arduino, Bluetooth, Website, Posyandu*

ABSTRAK

Kegiatan Posyandu merupakan hal penting yang harus dilakukan demi memastikan kesehatan balita dengan cara mengukur tinggi dan berat badan balita. Saat ini kegiatan Posyandu masih dilakukan dengan menggunakan alat ukur manual yang kadang membahayakan balita, dan cukup merepotkan bagi Kader Posyandu dalam melakukan pencatatan dan perekapan data. Perlu dilakukan perancangan produk yang dapat mendukung keberlangsungan kegiatan Posyandu sehingga dapat berjalan dengan nyaman dan aman. Pada penelitian ini akan dilakukan perancangan produk BY-PRO yang merupakan solusi utama dari permasalahan tersebut. Perancangan desain alat terdiri dari mikrokontroler arduino sebagai pengolah data, sensor load cell (berat), modul optocoupler (tinggi), dan sensor K02 (sensor sentuh) yang terhubung ke android melalui modul (bluetooth) hc-05. Aplikasi android dapat digunakan oleh kader posyandu untuk menampilkan data identitas dan pengukuran balita yang tersimpan ke database serta terdistribusi ke sistem informasi berbasis website untuk pemantauan kesehatan secara

realtime oleh bidan. Hasil rancangan tersebut kemudian dipaparkan dalam bentuk video sehingga mampu memberikan gambaran dari produk yang dirancang.

Kata Kunci— Pengukur Tinggi Badan Balita, Pengukur Berat Badan Balita, Android, Arduino, Bluetooth, Website, Posyandu

I. PENDAHULUAN

Pentingnya kesehatan seringkali diabaikan oleh masyarakat, ini menyebabkan rendahnya upaya tindakan preventif akan suatu penyakit. Padahal upaya pemantauan kesehatan yang dilakukan sejak anak masih didalam kandungan sampai 5 tahun pertama kehidupannya, ditujukan untuk mempertahankan kelangsungan hidupnya sekaligus meningkatkan kualitas hidup anak agar mencapai tumbuh kembang optimal [1]. Ditambah dengan mahalnya upaya untuk pengobatan suatu penyakit mendorong terbentuknya kegiatan Pos Pelayanan Terpadu (Posyandu). Posyandu merupakan salah satu bentuk upaya kesehatan bersumber daya masyarakat yang dikelola dan diselenggarakan dari, oleh, untuk, dan bersama masyarakat dalam penyelenggaraan pembangunan kesehatan, guna memberdayakan masyarakat dan memberi kemudahan masyarakat dalam memperoleh pelayanan kesehatan dasar, utamanya untuk mempercepat penurunan angka kematian ibu dan bayi [2][3].

Kementerian Kesehatan Republik Indonesia membuat program posyandu yaitu pengukuran berat dan tinggi badan bayi usia 1-5 tahun (balita) bulanan di posyandu dengan menggunakan Kartu Menuju Sehat (KMS). KMS memuat kurva pertumbuhan normal anak berdasarkan indeks *antropometri* berat badan menurut umur [4]. Untuk mewujudkan kota sehat diperlukan pelayanan kesehatan yang prima [5]. Sehingga semakin tinggi pelayanan kesehatan terhadap masyarakat diharapkan akan semakin meningkat derajat kesehatan masyarakat.

Dalam melaksanakan atau mendukung program bulanan yaitu posyandu, tentunya harus adanya partisipan yaitu masyarakat yang khususnya memiliki balita. Dalam program posyandu untuk balita setiap bulan, pengukuran berat dan tinggi badan adalah hal yang tidak dapat dikesampingkan. Di Indonesia banyak dipasarkan alat untuk mengukur berat badan dan tinggi badan pada bayi hanya saja alat ukurnya masih manual dan masih terpisah-pisah antara satu dengan yang lainnya. Dalam perekapan data juga masih

menggunakan alat tulis manual. Sebuah penelitian membahas tentang sistem yang dikembangkan generasi muda Indonesia teknologi tinggi badan dan berat badan bayi dibangun dengan *mikrokontroler* yang dapat menyelesaikan permasalahan ini yaitu sistem alat ukur berat badan, tinggi badan dan suhu badan di posyandu berbasis *android* [6].

Dari hal tersebut, kami memiliki gagasan untuk membuat rancangan sistem pengukuran berat badan, tinggi badan, dan perekapan data balita secara otomatis yang memudahkan kegiatan posyandu dan disajikan dalam bentuk video perancangan. Untuk membangun sistem pengukuran berat badan, tinggi badan, dan perekapan data balita secara otomatis ini, diperlukan beberapa komponen utama salah satunya adalah arduino. Arduino adalah sebuah *prototyping platform* berlisensi terbuka yang didasarkan pada kemudahan penggunaan perangkat keras dan perangkat lunak serta merupakan papan rangkaian elektronik yang bersifat *open source* dengan komponen utama sebuah *chip* mikrokontroler dengan jenis AVR dari perusahaan Atmel [7].

Rancangan sistem pengukuran ini juga dilengkapi dengan fitur *monitoring* tumbuh kembang balita melalui grafik yang disesuaikan dengan Kartu Menuju Sehat (KMS), serta dilengkapi dengan

adanya fitur penyimpanan data balita untuk memudahkan perekapan data bagi kader posyandu yang menggunakan antarmuka *Android. User Interface* atau tampilan antarmuka pengguna merupakan bagian yang penting dalam sebuah sistem atau aplikasi. Antarmuka adalah bagian sistem yang digunakan untuk berinteraksi langsung dengan pengguna [8]. Sehingga sistem informasi yang diusulkan dibuat dengan berbasis *website* dengan *user interface* yang menarik dan berfungsi untuk pemantauan kesehatan secara *realtime* oleh bidan dan bidan dapat mengajukan bantuan PMT (Pemberian Makanan Tambahan) untuk ibu dan anak.

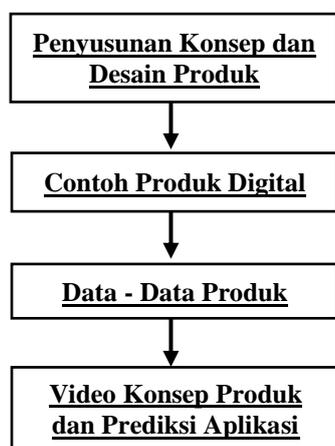
Rancangan perangkat ini didasarkan pada beberapa hal tersebut dengan berfokus pada pengukuran berat dan tinggi badan balita yang dilengkapi grafik pertumbuhan serta fitur rekap data bagi kader, dan menggunakan antarmuka *Android* sebagai sistem *monitoring* dan kontrol pada alat pengukuran berat dan tinggi badan balita secara otomatis.

Secara ringkas penelitian ini bertujuan untuk membuat sebuah rancangan perangkat guna mendukung kegiatan posyandu yang telah dimodifikasi dan dilengkapi dengan fitur-fitur yang memang dibutuhkan oleh Kader posyandu. Rancangan tersebut kemudian disajikan dalam bentuk video perancangan, untuk

memberikan gambaran dari inovasi yang akan diciptakan. Dengan dibuatnya video perancangan ini juga akan memberikan gambaran cara kerja dan penggunaan dari perangkat yang akan dikembangkan.

II. METODE PENELITIAN

Secara garis besar, metode yang digunakan dalam pembuatan video perancangan ini terdiri dari beberapa tahapan diantaranya Penyusunan Konsep dan Desain Produk, Contoh Produk Digital, Data-data Produk, serta Pembuatan Video Konsep Produk dan Prediksi Aplikasi. Tahapan metode tersebut ditunjukkan pada Gambar 1 dan selanjutnya akan dijelaskan pada sub-bab berikutnya.



Gambar 1. Gambar Metode Pelaksanaan

2.1 Penyusunan Konsep dan Desain Produk

Penyusunan konsep dan desain ini merupakan langkah awal membuat sketsa

rancangan awal dari desain alat pengukur berat dan tinggi badan balita serta *user interface* dari aplikasi *android* yang dilengkapi dengan sistem informasi berbasis *website*, penyusunan konsep dan desain ini berdasarkan :

a. Literatur dan Jurnal

Pada tahap ini mencari rujukan bagaimana membuat alat ukur berat dan tinggi badan balita serta perekapan data secara otomatis dengan menggunakan antarmuka *android* pada literatur maupun jurnal terdahulu. Beberapa penelitian terdahulu yang sejenis diantaranya yaitu: (i) Penelitian yang dilakukan oleh Heri Arbianto tahun 2018 [9] yang berjudul “Rancang Bangun Alat Pengukur Berat Badan dan Tinggi Badan Balita Dengan Metode *Antropometri* Berbasis *Arduino Uno*”. (ii) Penelitian berikutnya yang dilakukan oleh Cahyono dan Suprayitno tahun 2018 [6] yang berjudul “Alat Ukur Berat Badan, Tinggi Badan dan Suhu Badan Di *Posyandu* Berbasis *Android*”. Kedua penelitian tersebut menggunakan sensor ultrasonik untuk mengukur tinggi badan balita, dan menggunakan *load cell* pada pengukuran berat badan balita. Perangkat yang dibuat pada kedua

penelitian tersebut tidak dilengkapi dengan fitur penyimpanan data.

b. **Pengalaman Pribadi**

Selain mencari rujukan melalui literatur dan jurnal yaitu dengan pengalaman pribadi di lingkungan sekitar yang masih menggunakan alat ukur manual maupun otomatis dengan menampilkan hasil pengukuran di layar LCD.

Dengan adanya literatur dan jurnal serta pengalaman pribadi tersebut dalam membuat konsep desain alat dan aplikasi dapat tergambarkan dengan menambahkan fitur pendataan otomatis untuk kader posyandu.

2.2 Contoh Produk Digital

Pada tahap selanjutnya yaitu dengan melakukan penelusuran digital untuk mendapatkan contoh produk atau alat yang setara yang pernah diproduksi. Dari hasil penelusuran tersebut diperoleh

salah satu contoh produk yang pernah dibuat oleh kelompok Program Kemitraan Masyarakat yang beranggotakan Pauladie Susanto, S.Kom., M.T. , Dr. Susijanto Tri Rasmana, S.Kom., M.T. , dan Weny Indah Kusumawati, S.Kom., M.MT. dari STIKOM Surabaya yang dilaksanakan pada tahun 2018 [10] dengan membuat sistem informasi untuk memantau tumbuh kembang balita.

Contoh produk tersebut sudah diimplementasikan dan dapat dilihat pada <https://tk.dinamika.ac.id/post/73/implementasi-perangkat-pemantauan-tumbuh-kembang-balita-di-posyandu>.

2.3 Data – Data Produk

Dari beberapa contoh produk yang diperoleh beberapa peneliti terdahulu diatas akan diulas sebagai perhitungan efisiensi. Contoh produk yang diperoleh tentunya yang setara dengan produk yang akan direncanakan sebagai perbandingan. Beberapa hasil ulasan dari penelitian tersebut ditampilkan pada Tabel 1 dan dijelaskan pada bagian berikutnya.

Tabel 1. Hasil Ulasan Produk yang Setara Sebagai Bandingan

Peneliti	Alat Pengukur	Kelebihan	Kekurangan
Heri Abrianto (2018)	Pengukur Tinggi : Sensor Ultrasonik Pengukur Berat : Load Cell	- Tampil pada layar LCD (Liquid Crystal Display)	- Sensor ultrasonik mengandung frekuensi sinyal yang tidak baik untuk balita. - Tidak ada penyimpanan data secara otomatis.
Cahyono dan Suprayitno (2018)	Pengukur Tinggi : Sensor Ultrasonik Pengukur Berat : Load Cell	- Tampil pada layar LCD (Liquid Crystal Display) - Tersedia aplikasi android yang digunakan untuk menampilkan hasil pengukuran.	- Sensor ultrasonik mengandung frekuensi sinyal yang tidak baik untuk balita. - Tidak ada penyimpanan data secara otomatis.
Paulide Susanto, dkk (2018)	Pengukur Tinggi : Detektor Ketinggian Pengukur Berat : Timbangan Digital	- Hasil pengukuran tampil pada layar LCD (Liquid Crystal Display) - Tersedia sistem informasi berbasis desktop.	- Proses pengukuran berat badan masih dilakukan dengan timbangan digital secara manual.

Dari hasil ulasan data – data pada Tabel 1 dapat disimpulkan bahwa masih ada peluang yang dapat dikembangkan dari beberapa produk di atas. Pada BY-PRO untuk mengukur tinggi atau panjang pada balita menggunakan modul *optocoupler* berbeda dengan penelitian sebelumnya dimana menggunakan *sensor ultrasonic*. Modul *optocoupler* tidak mengandung frekuensi sinyal layaknya *sensor ultrasonic* sehingga lebih aman jika digunakan untuk mengukur tinggi badan balita, serta bisa dikendalikan secara otomatis dengan memasukkan kode

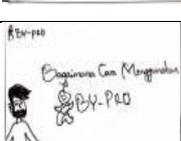
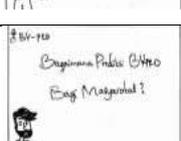
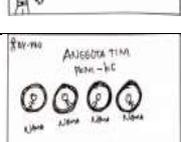
program pada *mikrokontroler*. Pada BY-PRO hasil pengukuran selain ditampilkan pada LCD juga ditampilkan pada sistem *android*, dimana sistem *android* tersebut dapat sekaligus mengontrol alat untuk melakukan pengukuran dan perekapan data balita secara otomatis untuk memantau tumbuh kembangnya balita. Data balita yang tersimpan dapat di cetak menjadi file pdf yang tidak bisa dirubah secara manual.

2.4 Video Konsep Produk dan Prediksi Aplikasi

Pada tahap ini dilakukan dengan konsep virtual dan digital dalam aplikasi

alat yang dihasilkan melalui video. Video tersebut berisikan animasi untuk merepresentasikan produk yang dirancang mulai dari desain, alur kerja, proses pembuatan, dan produk dalam visualisasi digital. Berikut ini adalah susunan gambar dari *storyboard*, yang ditunjukkan pada Tabel 2.

Tabel 2. *Storyboard*

No	Gambar	Keterangan	Durasi
1.		Menampilkan <i>opening</i> dengan warna latar belakang biru.	00.00 – 00.05
2.		Menampilkan situasi di Indonesia terkait Kesehatan balita yang stunting dan layanan posyandu	00.06 – 00.55
3.		Menampilkan video animasi 2D tentang solusi dan konsep alat dengan komponen yang muncul bergantian	00.56 – 01.45
4.		Menampilkan video animasi 2D tentang contoh produk yang setara sebagai perbandingan	01.46 – 02.35
5.		Menampilkan video animasi 2D tentang keunikan dan kebaharuan BY-PRO	02.36 – 03.55
6.		Menampilkan video animasi 2D tentang proses pembuatan BY-PRO dengan komponen alat muncul bergantian	03.56 – 05.00
7.		Menampilkan video animasi 2D tentang cara penggunaan BY-PRO dengan komponen alat muncul bergantian	05.01 – 06.02
8.		Menampilkan video animasi 2D tentang prediksi BY-PRO bagi masyarakat	06.03 – 06.54
9.		Menampilkan video animasi 2D foto tim dari PKM-KC	06.55 – 07.00

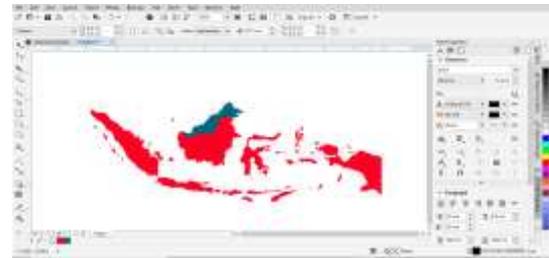
III. HASIL DAN PEMBAHASAN

Berdasarkan metode penelitian yang sudah dijelaskan pada bab sebelumnya, terciptalah rancangan produk BY-PRO yang dipaparkan melalui video perancangan untuk memberikan gambaran mengenai rancangan produk.

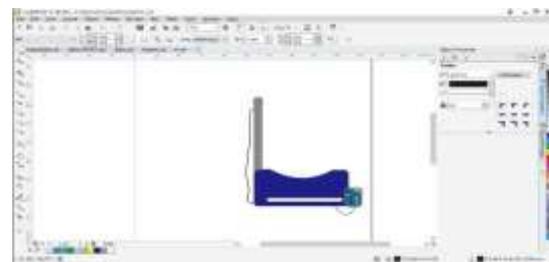
BY-PRO sendiri memiliki beberapa keunikan dan kebaruan diantaranya adalah pengukuran tinggi badan dilakukan secara otomatis dengan menggunakan modul *optocoupler* sehingga tidak ada radiasi yang dipancarkan. BY-PRO juga dilengkapi dengan aplikasi android yang terkoneksi melalui *bluetooth* untuk pencatatan oleh Kader posyandu. Serta dilengkapi dengan sistem informasi *website* yang dapat digunakan untuk pengelolaan kegiatan posyandu oleh Bidan, dan dapat melakukan pemantauan secara *realtime*. Kelebihan tersebut nantinya juga akan dijelaskan pada video rancangan yang dibuat dan disertai dengan rencana tahapan proses pembuatan BY-PRO sesuai dengan *storyboard* yang sudah dibuat.

Selanjutnya hasil pembuatan video dijelaskan bertahap mulai dari, hasil tampilan dan perancangan alat ini yang disajikan pada Gambar 2 (a) merupakan desain peta Indonesia terkait kesehatan

balita. Gambar 2 (b) tampilan desain alat pengukuran berat dan tinggi badan balita.



Gambar 2.a. Desain Peta Indonesia

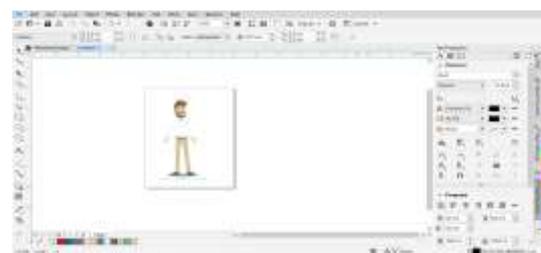


Gambar 2.b. Alat Pengukur Tinggi dan Berat Badan Balita

Setelah peta dan alat yang didesain selanjutnya pada Gambar 3 (a) desain sistem *android* dan Gambar 3 (b) desain karakter.



Gambar 3.a. Desain Sistem Android



Gambar 3.b. Desain Karakter

Setelah pembuatan desain selesai, selanjutnya tahap *editing* yang merupakan pengeditan video dari mulai penggabungan, pewarnaan, hingga menambahkan transisi dalam video. Berikut pada Gambar 4 (a) merupakan pembukaan video dengan warna permukaan biru. Gambar 4 (b) menggambarkan kesehatan balita di Indonesia.

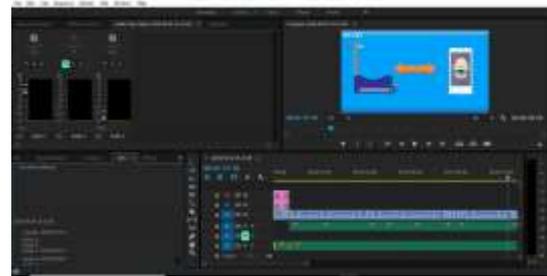


Gambar 4.a. Pembukaan Video



Gambar 4.b. Keadaan Kesehatan Balita di Indonesia

Selanjutnya Gambar 5(a) penggabungan bagian solusi dan konsep dari alat. Gambar 5 (b) contoh alat yang sudah ada atau setara.

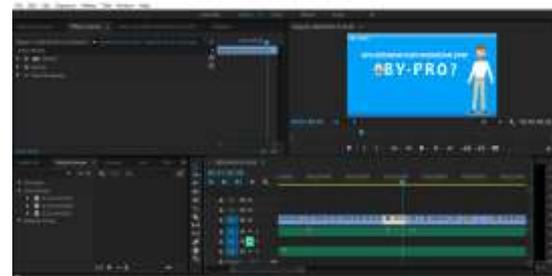


Gambar 5.a. Solusi dan Konsep Alat



Gambar 5.b. Contoh Alat Yang Sudah Ada atau Setara

Dari contoh yang sudah ada tentunya BY-PRO memiliki keunikan dan kebaruan serta proses pembuatan secara garis besar. Gambar 6 (a) merupakan bagian dari keunikan BY-PRO. Gambar 6 (b) proses pembuatan BY-PRO.



Gambar 6.a. Keunikan BY-PRO



Gambar 6.b. Proses Pembuatan BY-PRO

Setelah penjelasan proses pembuatan dilanjutkan dengan cara menggunakan BY-PRO serta kreativitas dan manfaat produk yang dihasilkan bagi masyarakat. Gambar 7 (a) cara menggunakan BY-PRO. Gambar 7 (b) kreativitas dan manfaat bagi masyarakat.



Gambar 7.a. Cara Menggunakan BY-PRO



Gambar 7.b. Kreativitas Untuk Masyarakat

Untuk selanjutnya yaitu bagian pengenalan anggota tim yang ditunjukkan pada Gambar 8 serta setelah selesai dalam proses *editing* pada Gambar 9 yaitu tampilan pemutar video alur perwujudan produk BY-PRO yang dibuat sesuai dengan daya yang didapat dari literatur, jurnal, dan pengalaman pribadi.



Gambar 8. Anggota Tim



Gambar 9. Tampilan Pemutar Video

Dalam pembuatan video alur perwujudan produk BY-PRO peneliti menggunakan data peneliti terdahulu sehingga dapat merancang alat tersebut dengan ditambahkan beberapa pembaruan. Selanjutnya, pengujian BY-PRO dilakukan dengan membandingkan hasil pengukuran manual yang dilakukan oleh Kader posyandu dengan hasil pengukuran yang diperoleh menggunakan BY-PRO. Dari hasil perbandingan tersebut akan diperoleh tingkat kesesuaian dari produk yang dibuat dengan pengukuran secara manual.

IV. SIMPULAN

Berdasarkan hasil dari video alur perwujudan produk BY-PRO diatas, didapat kesimpulan sebagai berikut :

1. Alat ukur tinggi badan, berat badan, dan pendataan balita berbasis arduino dengan antar muka *android*.
2. Dilengkapi dengan sistem informasi berbasis *website*.
3. Diharapkan mampu menciptakan program bulanan posyandu lebih nyaman, aman, efisien, mudah, dan dapat digunakan dalam keadaan normal atau ketika dalam keadaan pandemic seperti COVID-19.

DAFTAR PUSTAKA

- [1] F. A. K, I. A. Hamsah, D. Darmiati, and M. Mirnawati, "Deteksi Dini Tumbuh Kembang Balita di Posyandu," *J. Ilm. Kesehat. Sandi Husada*, vol. 12, no. 2, pp. 1003–1008, 2020, doi: 10.35816/jiskh.v12i2.441.
- [2] Kementerian Kesehatan, *Pedoman umum pengolahan posyandu*. Jakarta: Menteri Kesehatan RI, 2011.
- [3] E. Saepuddin, E. Rizal, and A. Rusmana, "Peran Posyandu Sebagai Pusat Informasi Kesehatan Ibu dan Anak Posyandu," *Rec. Libr. J.*, vol. 3, no. 2, pp. 201–208, 2017, doi: 10.20473/rlj.v3-i2.2017.201-208.
- [4] F. Febry, *Pemantauan Pertumbuhan Balita di Posyandu*, vol. 3. 2012.
- [5] M. R. Napirah, A. Rahman, and A. Tony, "Faktor-Faktor Yang Berhubungan Dengan Pemanfaatan Pelayanan Kesehatan Di Wilayah Kerja Puskesmas Tambarana Kecamatan Poso Pesisir Utara Kabupaten Poso," *J. Pengemb. Kota*, vol. 4, no. 1, p. 29, 2016, doi: 10.14710/jpk.4.1.29-39.
- [6] T. H. A. Cahyono and E. A. Suprayitno, "Alat Ukur Berat Badan, Tinggi Badan dan Suhu Badan di Posyandu Berbasis Android," *ELINVO (Electronics, Informatics Vocat. Educ.*, vol. 3, no. 1, pp. 31–38, 2018, doi: 10.21831/elinvo.v3i1.20221.
- [7] S. Yatmono, "Pengembangan Aplikasi User Interface Android untk Pengukur Jarak berbasis Arduino dan Bluetooth," *J. Edukasi Elektro*, vol. 1, no. 2, pp. 134–138, 2017.
- [8] M. Nauval *et al.*, "Analisis Komponen Desain Layout , Warna , dan Kontrol Pada Antarmuka Pengguna Aplikasi Mobile Berdasarkan Kemudahan Penggunaan (Studi Kasus : Aplikasi Olride)," vol. 7, no. 1, 2018.
- [9] H. Abrianto, "Rancang Bangun Alat Pengukur Berat Badan dan Tinggi

- Badan Balita dengan Metode Antropometri Berbasis Arduino Uno,” *Fakultas Sains dan Teknologi*. Universitas Islam Negeri Alaudin Makassar, Makassar, 2018, doi:
10.1051/mateconf/201712107005.
- [10] “Implementasi Perangkat Pemantauan Tumbuh Kembang Balita Di Posyandu,” *Universitas Dinamika*, 2018.
<https://tk.dinamika.ac.id/post/73/implementasi-perangkat-pemantauan-tumbuh-kembang-balita-di-posyandu>.

SISTEM PAKAR FORWARD CHAINING , FUZZY-MAX DAN CERTAINTY FACTOR AYAM PEDAGING

Asep Afandi¹, Dwi Marisa Efendi²

¹ Sistem Informasi, ² Teknologi Komputer , STMIK DIAN CIPTA CENDIKIA KOTABUMI
Jl. Raya Candimas No. 03, Lampung Utara - Indonesia 35418
Telp. (0721) 787214 Fax. (0721) 700261
e-mail: Asepfandi189@Gmail.Com, Dwimaris89@Gmail.Com

ABSTRACT

Chicken farming is one of the promising business potentials, but in management and care, it is very important to determine the success of chicken farming. Often in the care of negligent employees so that they are exposed to disease outbreaks. With various types of diseases that attack the symptoms are very similar and even the same as other diseases, therefore an expert system can be used to diagnose a disease by looking at the characteristics of the disease suffered, and how the solution is to treat or prevent the disease. In the study, it discusses 8 types of broiler diseases, where the expert system method used is the Fuzzy Max method, Forward Chaining, and Certainty Factor. From the results of the Fuzzy Max method, the results showed an accuracy of 80% - 90% for all types of diseases, while the Certainty Factor method showed 96% - 99% for all types of diseases.

Keywords— *Expert System, Fuzzy Max, Certainty Factor, Forward Chaining.*

ABSTRAKSI

Peternakan ayam salah satu potensi bisnis yang menjanjikan, akan tetapi dalam manajemen, dan perawatan sangat menentukan keberhasilan peternakan ayam. Sering sekali dalam perawatan pegawai lalai sehingga terkena wabah penyakit. Dengan Berbagai Jenis penyakit yang menyerang gejalanya sangatlah mirip bahkan sama dengan penyakit lainnya oleh karena itu sistem pakar dapat digunakan mendiagnosa suatu penyakit dengan melihat ciri-ciri penyakit yang diderita, dan bagaimana solusi dalam mengobati atau mencegah penyakit tersebut. Dalam penelitian membahas tentang 8 jenis penyakit Broiler, dimana metode system pakar yang digunakan adalah metode *Fuzzy Max, Forward Chaining, dan Certainty Factor*. Dari hasil penelitian Metode *Fuzzy Max* didapat hasil menunjukkan keakuratan 80% - 90% untuk semua jenis penyakit sedangkan metode *Certainty Factor* menunjukkan 96%- 99% untuk semua jenis penyakit.

Kata Kunci—Sistem pakar , Fuzzy Max, Certainty Factor, Forward Chaining.

I. PENDAHULUAN

Peternakan ayam salah satu potensi bisnis yang menjanjikan, akan tetapi dalam manajemen, dan perawatan sangat menentukan keberhasilan peternakan ayam. Sangat seing sekali dalam perawatan pegawai lalai sehingga terkena wabah penyakit. Dengan Berbagai Jenis penyakit yang menyerang gejalanya sangatlah mirip bahkan sama dengan penyakit lainnya oleh karena itu sistem pakar sangat dibutuhkan untuk mengetahui atau mendiagnosa suatu penyakit ayam pedaging dengan melihat ciri-ciri penyakit yang diderita dan solusi dalam mengobati atau mencegah penyakit tersebut. Dalam penelitian membahas 8 jenis penyakit ayam pedaging (Broiler), dimana metode system pakar yang digunakan adalah metode *Fuzzy Max*, *Forward Chaining* dan *Certainty Factor*.

Peneliti ini pengambilan data dibantu oleh seorang Petugas Penyuluh Lapangan dari PT. Ciomas Adisatwa cabang Lampung Tengah. Dimana data berupa data penyakit, gejala-gejala setiap penyakit dan obat-obatan didalamnya terdapat penjelasan tentang penyakit yang dapat disembuhkan, dan hal-hal pendukung lainnya. penulis menganalisa

apa saja yang dibutuhkan oleh user dalam sistem yang akan dibangun.

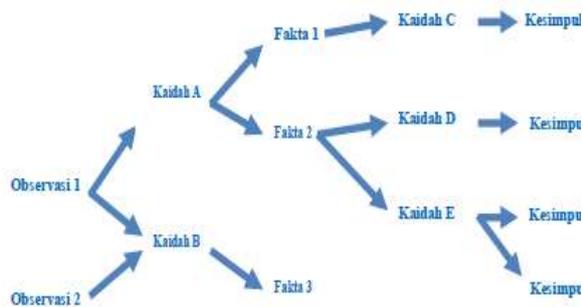
II. METODE PENELITIAN

2.1. Sistem Pakar

Sistem Pakar adalah aplikasi komputer Artificial Intelligence (AI) yang berisi basis pengetahuan dan mesin inferensi sebagai komponen utama[1]. Expert System adalah sistem komputer yang meniru, atau bertindak dalam segala hal, dengan kemampuan pengambilan keputusan seorang pakar manusia. Komponen utamanya adalah: Basis pengetahuan, dapat diperoleh dari buku, majalah, orang berpengetahuan, dll. Mesin inferensi, menarik kesimpulan dari basis pengetahuan.[2]

2.2. Forward Chaining

Sistem pakar *Forward chaining* yang dapat membuat kesimpulan tentang fakta yang ada menggunakan aturan, objek dan mengambil tindakan yang sesuai sebagai hasilnya[1]. Forward Chaining penalaran yang dimulai dari sekumpulan data menuju kesimpulan[3].



Gambar 1. Forward Chaining[3]

2.3. Fuzzy Mamdani

Bentuk distribusi yang memungkinkan keluaran prosedur defuzzifikasi cepat dan sederhana simple prosedurnya adalah mengambil salah satu nilai terbesar yang memiliki a derajat kebenaran maksimum (metode maxima)[4]. Pilihan yang mungkin adalah yang pertama (terkecil), yang terakhir (terbesar) atau, dalam kasus distribusi kemungkinan unimodal, nilai median. Sejauh ini metode maxima yang paling umum adalah memilih nilai rata-rata elemen dengan derajat kebenaran maksimum (metode MOM)[4].

2.4. Certainty Factor

Metode *Certainty Factor* adalah pilihan yang tepat dalam penelitian ini, karena pada dasarnya metode CF diasumsikan sebagai tingkat kepercayaan pakar terhadap data yang digunakan. Metode *Certainty Factor* hanya dapat memproses dua bobot dalam satu perhitungan. Untuk

bobot lebih dari 2 banyak, untuk melakukan perhitungan untuk menghindari masalah ketika bobot dihitung secara acak berarti tidak ada aturan untuk menggabungkan bobot karena kombinasi dari setiap hasil tersebut akan tetap sama. Penelitian sebelumnya menjelaskan hasil pencarian sistem pakar menunjukkan bahwa setiap gejala yang dipilih oleh pengguna akan dicari semua jenis penyakit yang memenuhi gejala tersebut[5]. Certainty Factor (CF) menunjukkan ukuran kepastian terhadap suatu fakta atau aturan. Metode faktor kepastian rumus umum sebagai berikut:

$$CFs[h,e] = MB[h,e] - MD[h,e]$$

Keterangan:

$CFs[h,e]$ = Faktor kepastian

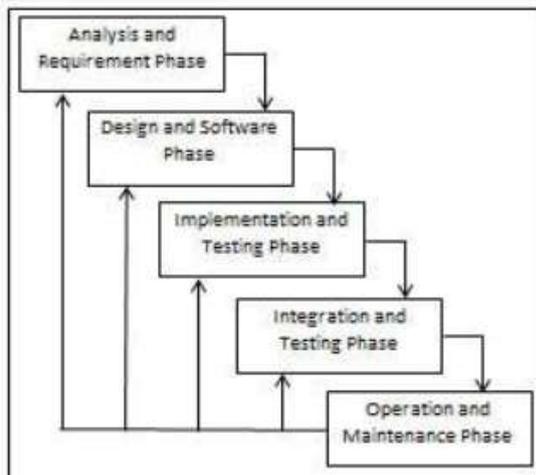
$MB[h,e]$ = Ukuran kepercayaan atau tingkat keyakinan terhadap hipotesis h, jika diberikan evidence e (antara 0 dan 1)

$MD[h,e]$ = Ukuran ketidakpercayaan atau tingkat keyakinan terhadap hipotesis h, jika diberikan evidence e (antara 0 dan 1)

2.5. Metode Penelitian

Pada Penelitian ini menggunakan metode pengembangan sistem *Waterfall*, dimana strategi pengembangan perangkat lunak pertama yang diterapkan, menyerupai desain yang digunakan di industri lain. Strategi ini memungkinkan proyek dipecah

menjadi beberapa fase[6]. Selain itu, dikenal sebagai siklus hidup klasik di mana metodologi ini adalah pendekatan siklus hidup paling dasar dan masih banyak digunakan dalam perangkat lunak. di bawah ini menunjukkan fase-fase pengembangan metodologi *Waterfall* [7].



Gambar 2. Ilustrasi Model *Waterfall* [8]

2.6. Knowledge Base

Knowledge Base Berisikan fakta, aturan, dan objek. Sistem berbasis pengetahuan tujuannya adalah untuk menentukan aturan dalam format yang intuitif dan mudah dipahami, ditinjau, dan bahkan diedit oleh pakar domain seorang pakar TI.[4]

Tabel 1. Penyakit Ayam Pedaging

Kode Penyakit	Nama Penyakit
P1	<i>Chronic Respiratory Disease (CDR)</i> (Ngorok)
P2	<i>Pullorum Disease</i> (Berak Kapur)
P3	<i>Gumboro Disease</i> (Gumboro)
P4	<i>Newcastle Disease</i> (Tetelo)
P5	<i>Infectious Coryza</i> (Snot)
P6	<i>Fowl Typhoid</i> (Tifus Ayam)
P7	<i>Coccidiosis</i> (Berak Darah)
P8	<i>Fowl kolera</i> (penyakit kolera)

Tabel 2. Gejala

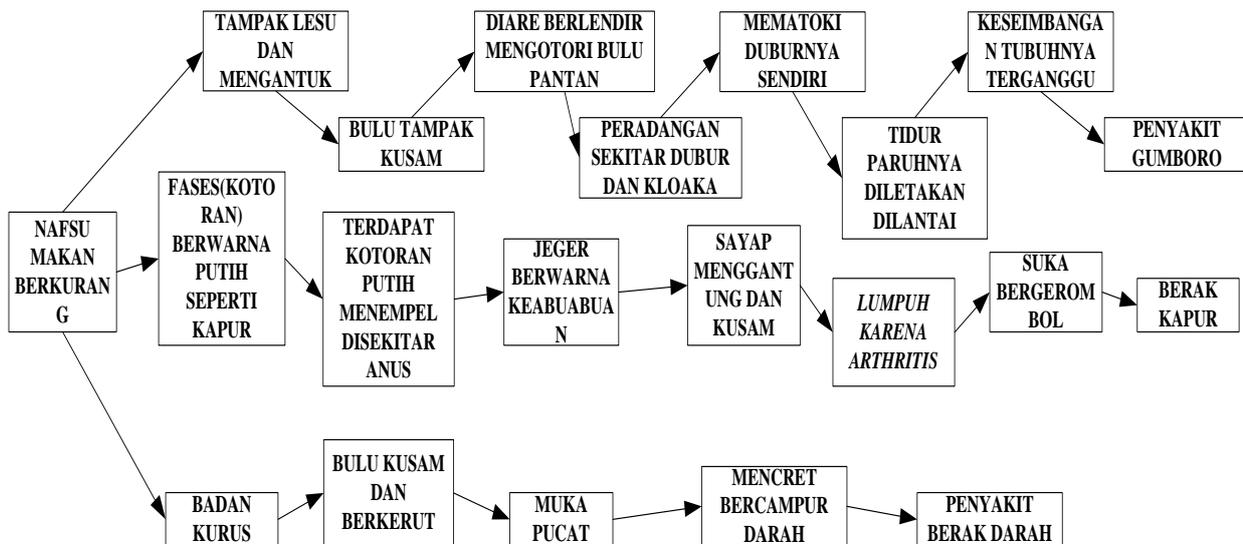
Kode Gejala	Nama Gejala
1	Mata terpejam seperti mengantuk
2	Keluar cairan berbusa dari mata
3	Jengger membengkak merah
4	Keluar nanah dari mata dan bau
5	keseimbangan tubuhnya terganggu
6	Sayap menggantung dan kusam
7	Terdapat kotoran putih menempel disekitas anus
8	Sayap terkulai
9	Tidur paruhnya diletakan dilantai
10	Duduk dengan sikap membungkuk

Tabel 3. Rules

Kode Gejala	Kode Penyakit							
	P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8
01						v		v
02				v	v			v
03				v	v			
04			v					
05		v						
06	v							
07	v							
08								
09	v			v	v		v	
10		v	v					v
11	v							
12								v
13				v				
14				v				
15				v				v
16						v		
17							v	
18		v						
19							v	
20			v					

2.7. Forward Chaining

Lapangan, PT. Ciomas Adisatwa, Lampung Tengah) maka didapat data



Gambar 3. Diagram Forward Chaining Penyakit Gumboro, Penyakit Berak Kapur, dan Penyakit Berak Darah.

seperti Tabel 5 dibawah ini:

Table 5. Interpretasi Nilai Bobot

2.8. Interpretasi Nilai Bobot

Jawaban pengguna terhadap pertanyaan diagnosis yang akan diolah menjadi sebuah nilai CF. Dari CF tersebut akan dihitung nilai CF Rule Gejala dan Penyakit dinamakan proses konversi sebuah nilai. Dibawah ini merupakan nilai evidence yang ditentukan oleh pakar[9].

Tabel 4. Nilai Bobot

Istilah	Bobot
Kurang Berpengaruh	0,1 s/d 0,4
Berpengaruh	0,5 s/d 0,7
Sangat Berpengaruh	0,8 s/d 1,0

Dari hasil Interview dan Observasi dengan seorang Pakar (Petugas Penyuluh

Nama Penyakit	Nama Gejala (Kode Gejala)	Interpretasi	Bobot
Berak Kapur (Pulloverum Disease)	Napsu makan menurun (09)	Kurang berpengaruh	0,3
	Feses (kotoran) berwarna putih seperti kapur (31)	Berpengaruh	0,7
	Terdapat kotoran putih menempel disekitar anus (37)	Sangat berpengaruh	0,8
	Jengger berwarna keabuan (48)	Berpengaruh	0,6
	Sayap menggantung dan kusam (06)	Kurang berpengaruh	0,3
	Lumpuh karena arthritis (peradangan persendian) (18)	Berpengaruh	0,5
	Suka bergerombol (23)	Kurang berpengaruh	0,4

Gumbo- bo- ro(Gum boro Dis- ease)	Nafsu makan berkurang (09)	Kurang ber- pengaru h	0,3
	Ayam tampak lesu dan mengantuk (04)	Kurang ber- pengaru h	0,3
	Bulu tampak kusam (20)	Kurang ber- pengaru h	0,3
	Diare berlendir yang men- gotori bulu pantat (26)	Ber- pengaru h	0,5
	Peradangan di sekitar dubur dan kloaka (33)	Sangat ber- pengaru h	0,8
	Mematoki duburnya sendiri (34)	Ber- pengaru h	0,6
	Tidur paruhnya diletakan di lantai (39)	Ber- pengaru h	0,5
	Keseimbangan tubuhnya terganggu (35)	Ber- pengaru h	0,5
Berak Darah(Coc- cidosis)	Nafsu makan berkurang (09)	Kurang ber- pengaru h	0,3
	Badan kurus (07)	Kurang ber- pengaru h	0,3
	Bulu kusam dan berkerut (08)	Ber- pengaru h	0,5
	Muka pucat (19)	Ber- pengaru h	0,5
	Mencret bercampur darah (17)	Sangat ber- pengaru h	0,9

2.9. Certainty Factor

Berdasarkan data dari tabel 7 didapat

:

Penyakit Berak Darah :

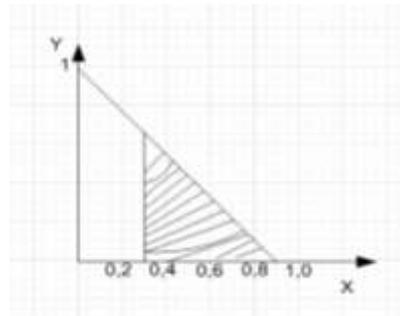
$$P007 = MB(G,17) + (MB(G,17) * (1 - MB(G,17)))$$

$$P007 = 0.90 + (0.90 * (1-0.90)) = 0.99$$

2.10. Fuzzy Mamdani Max

1. Penyakit Berak Darah (*Coccidiosis*)

Berdasarkan data dari table 7 maka didapat seperti gambar grafik 4 dibawah ini:



Gambar 4. Grafik Penyakit Berak Darah

Berdasarkan Data dari Gambar maka

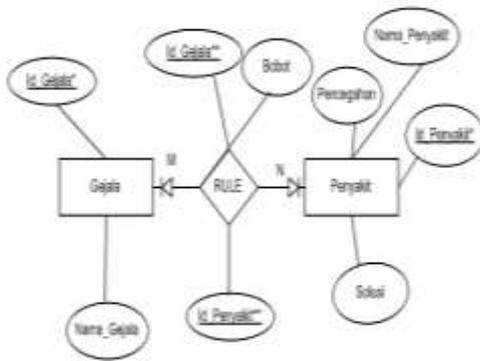
didapat:

$$\begin{aligned} \mu_{df}(xi) &= \text{Max}(\mu_{df}(xi), \mu_{kf}(xi)) \\ &= \text{Max}(0,5;0,9) \\ &= \text{Max}(0,9) \end{aligned}$$

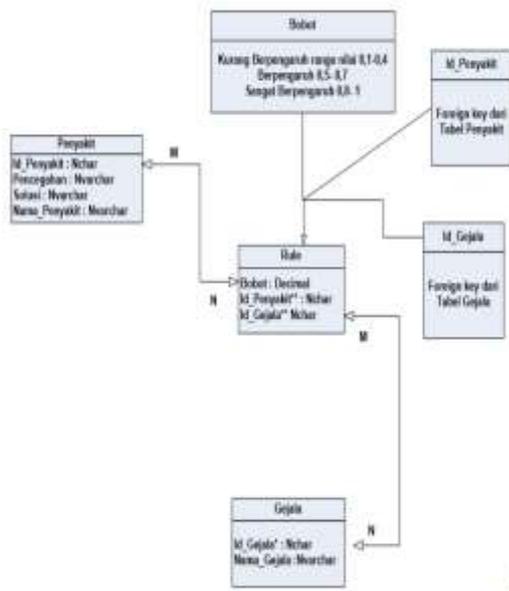
2.11. Perancangan Sistem Pakar

2.11.1. Database

Pada tahapan ini menjelaskan bagaimana proses terbentuknya suatu database, yang terdiri dari Entity Relationship Diagram (ERD), Class Diagram, dan Struktur Basis Data.



Gambar 5. Entity Relationship Diagram



Gambar 6. Class Diagram

Class diagram struktur aplikasi detail untuk database.

III. Hasil dan pembahasan

Dari uraian pada metode penelitian bahwa bobot gejala yang nilainya diatas 0,6 dijadikan Gejala Primary dan Scondary dimana nilai bobot kedua terbesar untuk metode *Certainty Factor* sedangkan nilai yang paling besar sebagai

nilai maksimumkan dijadikan nilai banding dengan rata-rata untuk metode *Fuzzy Max* yang diterapkan dalam aplikasi sistem pakar sehingga didapat:



Gambar 7. Pemilihan Gejala Pada Pertanyaan Diagnosa.



Gambar 8. Hasil Diagnosa.

Hasil implementasi aplikasi program sistem pakar penyakit ayam pedaging, seperti pada Tabel 6 :

Tabel 6. Persentase Hasil program

N	Name of disease	Certain	Fuzz
---	-----------------	---------	------

o.		y Factor (%)	y Max (%)
1	Snoring Disease (Chronic Respiratory Disease (CDR))	96%	80%
2	Chalk Defecation (Pullorum Disease)	96%	80%
3	Gumboro Disease	96%	80%
4	Tetelo disease (NewCastle Disease)	99%	90%
5	Snot's disease (Coryza)	98%	85%
6	Chicken Typhus (Fowl Typhoid)	96%	80%
7	Intestinal tract disease (Coccidiosis)	99%	90%
8	Chicken Cholera Disease (Fowl Cholera)	96%	80%

IV. KESIMPULAN

Dari hasil penelitian bahwa metode *Forward Chaining* digunakan untuk pengambilan kesimpulan, sedangkan Metode *Fuzzy Max* didapat hasil menunjukkan keakuratan 80% - 90% untuk semua jenis penyakit dan *Certainty Factor* menunjukkan 96% - 99% untuk semua jenis penyakit.

V. DAFTAR PUSTAKA

- [1] I. A. Alshawwa, M. Elkahlout, H. Q. El-mashharawi, and S. S. Abu-naser, "An Expert System for Depression Diagnosis," *Int. J. Acad. Heal. Med. Res.*, vol. 3, no. 4, pp. 20–27, 2019, [Online]. Available: www.ijeaais.org/ijjahmr.
- [2] F. M. Salman and S. Abu-Naser, "Expert System for Castor Diseases and Diagnosis," *Int. J. Eng. Inf. Syst.*, vol. 3, no. 3, pp. 1–10, 2019.
- [3] B. Herawan Hayadi, A. Bastian, K. Rukun, N. Jalinus, Y. Lizar, and A. Guci, "Expert system in the application of learning models with Forward Chaining Method," *Int. J. Eng. Technol.*, vol. 7, no. 2.29 Special Issue 29, pp. 845–848, 2018, doi: 10.14419/ijet.v7i2.29.14269.
- [4] D. T. Pham and M. Castellani, "Action aggregation and defuzzification in Mamdani-type fuzzy systems," *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 216, no. 7, pp. 747–759, 2002, doi: 10.1243/09544060260128797.
- [5] W. U. Setiabudi, E. Sugiharti, and F. Y. Arini, "Expert System Diagnosis Dental Disease Using Certainty Factor Method," *Sci. J. Informatics*, vol. 4, no. 1, pp. 43–50, 2017, doi: 10.15294/sji.v4i1.8463.
- [6] B.-A. Andrei, A. Casu-Pop, S.-C. Gheorghe, and C.-A. Boiangiu, "A Study on Using Waterfall and Agile Methods in Software Project Management," *J. Inf. Syst. Oper. Manag.*, no. June, pp. 125–135, 2019, [Online]. Available: <http://0->

- search.proquest.com.ditlib.dit.ie.tud
ublin.idm.oclc.org/docview/223782
8314?accountid=10594.
- [7] N. F. Basir *et al.*, “Sweet8Bakery
Booking System,” *Acta Electron.
Malaysia*, vol. 2, no. 2, pp. 14–19,
2018, doi:
10.26480/aem.02.2018.14.19.
- [8] A. Details, “Acta Electronica
Malaysia (AEM),” vol. 2, no. 2, pp.
10–13, 2018.
- [9] A. Afandi, “Sistem Pakar
Identifikasi Penyakit Ayam
Pedaging,” *J. Cendikia*, vol. 16, no.
1, pp. 58–67, 2018, [Online].
Available:
[https://jurnal.dcc.ac.id/index.php/J
C/article/view/65](https://jurnal.dcc.ac.id/index.php/JC/article/view/65).

PENGELOLAAN DATA PRODUKSI BERBASIS WEB PADA CV. SUMBER PROTEINA BANDAR LAMPUNG

TM Zaini¹, Driya Wiryawan², Aryadi³

^{1,3} Fakultas Ilmu Komputer, Institut Informatika & Bisnis Darmajaya
Jl. Z.A Pagar Alam No 93, Bandar Lampung - Indonesia 35142
Telp. (0721) 787214 Fax. (0721)700261.

² Fakultas Ekonomi dan Bisnis Universitas Lampung
Jl. Prof. Brojonegoro No.1. Gedung Meneng Bandar Lampung – Indonesia 35141
e-mail : tmzaini@ darmajaya.ac.id, driyawiryawan@gmail.com
aryadi_97e@yahoo.com

ABSTRACT

CV. Sumber Proteina is a company that engaged in animal husbandry that produces and distributes chicken eggs. The reporting process at this company is still done manually, so it takes quite a long time to be able to provide reporting data until it reaches the head office. Thus the report need long time to reach the manager. Based on this, this company need a website that can facilitate the delivery of reports and data processing reports, so as to minimize the time of sending data and managing data reports. Management of production data at this company can help decision-making parties at the management level to speed up the process of reporting and processing production data. The system development uses the RAD (Rapid Application Development) model and the design tool is the Unified Modeling Language (UML). The stages for developing RAD system begin at the planning requirements stage, the user design, the Rapid Construction stage and the cutover or system testing stage. The results of this study are in the form of a system that can provide information on the availability of chicken eggs and can also be used online for consumers.

Keywords—*Data Processing, Production, Report, Website.*

ABSTRAK

Perusahaan CV. Sumber Proteina adalah perusahaan yang bergerak dibidang peternakan yang memproduksi sekaligus sebagai distributor telur ayam. Proses pelaporan pada perusahaan ini masih dilakukan secara manual sehingga membutuhkan waktu yang cukup lama untuk dapat memberikan data pelaporannya hingga sampai di kantor pusat. Dengan demikian mengakibatkan lambatnya informasi untuk sampai ke *manager*. Berdasarkan hal tersebut maka dibutuhkan sebuah *website* yang dapat memudahkan pengiriman laporan dan pengolahan data laporan, sehingga dapat meminimalisir waktu pengiriman data dan pengelolaan data laporan. Pengelolaan data produksi pada perusahaan ini dapat membantu pihak pengambilan keputusan di tingkatan manajemen guna mempercepat proses pelaporan dan pengolahan data produksi. Pengembangan sistem menggunakan model RAD (*Rapid Application*

Development) dan sebagai *tool* perancangannya adalah *Unified Modelling Language (UML)*. Tahapan untuk pengembangan sistem RAD ini diawali dengan tahap *planning requirement*, tahap *user design* dan *workshop*, tahap *Rapid Construction* dan *cutover* atau tahap pengujian sistem. Hasil penelitian ini berupa sistem yang dapat memberikan informasi data ketersediaan telur ayam sekaligus dapat dimanfaatkan bagi konsumen secara online.

Kata kunci—*Pengolahan Data, Produksi, Pelaporan, Website.*

I PENDAHULUAN

Perkembangan teknologi saat ini telah berdampak terhadap pekerjaan, karena mempermudah suatu pekerjaan. Keberadaan teknologi yang berkembang dengan cepatnya adalah teknologi komunikasi yang meliputi hardware seperti komputer dan peralatan elektronik yang semakin banyak diminati oleh pengguna dimana pengembang dapat lebih menekankan pada studi kelayakan dan definisi sistem. [1] Teknologi Informasi telah merambah dan diterapkan di berbagai bidang antara lain bidang produksi dan peternakan. Kebutuhan akan pengolahan data dan pelaporan, merupakan masalah yang penting dalam manajemen sebuah perusahaan. Salah satunya adalah CV. Sumber Proteina yang merupakan suatu perusahaan yang bergerak dibidang peternakan dengan memproduksi telur ayam, pendistribusian telur ayam dan pembesaran ayam.

Adapun pengiriman laporan dilakukan oleh seorang kurir telur dan kemudian melakukan pengiriman pada waktu pagi hari di mana telah di kumpulkan telur-telur yang siap untuk dikirimkan, sedangkan keterlambatan proses laporan pada produksi telur untuk dapat disampaikan pada kantor pusat serta terhambatnya informasi untuk sampai kepada seorang *manager*, maka untuk itu diperlukan pemecahan masalah guna mempercepat pengelolaan data produksi dan data pelaporan.

Solusi yang ditawarkan adalah mengembangkan, Pengelolaan Data Produksi pada CV. Sumber Proteina ini yaitu kiranya dapat membantu pihak manajemen perusahaan untuk dapat mempercepat proses pelaporan data produksi dan pengolahan data produksi. Dengan adanya pengelolaan data produksi ini selanjutnya akan memudahkan proses pengolahan data dan pelaporan. Sebagai contoh, karyawan CV. Sumber Proteina hanya perlu *login*

ke *website* selanjutnya melakukan *input* data produksi, kemudian *website* akan mengolah data produksi, begitupun *manager* hanya perlu *login* ke *website* untuk melihat laporan produksi.

Website ini dapat terjemahkan sebagai suatu kumpulan halaman yang menampilkan informasi data teks, gambar diam atau bergerak, data animasi, suara, *video* serta gabungannya, baik yang bersifat tetap maupun dinamis yang membentuk suatu kesatuan bangunan yang saling terhubung dimana setiap unit disatukan dengan jaringan halaman (*hyperlink*) [2]

Informasi merupakan data dihasilkan dari proses pembentukan yang lebih bermanfaat dan lebih berarti bagi penggunaannya.[3]

Proses pengolahan basis data yang terdistribusi merupakan suatu proses dimana kegiatan aktivitas transaksi, pengambilan dan pemuktahiran data yang terjadi telah melalui dua atau lebih perangkat komputer yang umumnya terpisah letaknya dan tidak saling terhubung. [4]

II METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah pengembangan

aplikasi menggunakan metode *Rapid Application Development (RAD)*., yang dilakukan dengan menerapkan empat tahapan untuk pengembangan sistem informasi CV. Sumber Proteina, yaitu:

1. *Planning Requirement.*

Pada tahap ini dilakukan pengumpulan data-data dan informasi mengenai gambaran umum, wawancara dan identifikasi tugas pokok pada bagian produksi CV. Sumber Proteina Bandar Lampung.

2. *User Design.*

Pada tahapan ini dilakukan proses desain pada sistem yang di kembangkan dan kebutuhan sistem serta usulan pada sistem berbasis web. untuk CV. Sumber Proteina Bandar Lampung.

Adapun penggunaan tahapan analisis dan desain *Object Oriented Programming* [5] yaitu: (a). Komunikasi Pada tahap komunikasi mengestimasi kebutuhan sistem baru yang akan dikembangkan. Analisis di mulai dari pengumpulan data terkait produksi pada CV. Sumber Proteina Bandar Lampung, yang dilakukan dengan beberapa metode pengumpulan data yang telah dijabarkan diatas dan juga melakukan konsultasi dengan pelanggan/*stakeholder*. (b) Perencanaan secara cepat (*quick plan*). Tahap ini lebih menekankan pada tahap

desain sistem secara cepat, desain sistem dilakukan untuk menindak lanjuti tahap sebelumnya dan sebagai acuan pembuatan program. Pada fase ini penulis melakukan perancangan arsitektur menggunakan *Unified Modeling Language* (UML) dimana bahasa pemodelan ini merupakan sistem yang menerapkan konsep pemrograman berorientasi pada objek yang melakukan analisa serta menjabarkan secara detail yang diperlukan sistem.[6] Rancangan yang dibuat berupa diagram sebagai berikut: *Document Flowchart, UseCase, Class Diagram* dan *Sequence Diagram*. (c) Pembuatan Model (*Modeling Quick Design*), Pada tahap ini mentranslasikan dari desain yang telah dirancang ke pemodelan (*modeling*) dimana dilakukan pembuatan desain sistem (*Design User Interface*). (d) Konstruksi Prototipe (*Construction of Prototype*), Tahapan ini adalah tahapan yang membangun sistem berdasarkan desain sistem (*Design User Interface*) sebelumnya dimana pada tahapan ini dilakukan pengkodean-pengkodean yang menggunakan bahasa pemrograman. (e) Penyerahan sistem (*Deployment Delivery & Feedback*). Tahap ini adalah merupakan tahapan serah terima dari sistem yang telah dibuat kepada CV. Sumber Proteina Bandar

Lampung. Umpan balik akan terjadi saat perawatan sistem, umpan balik berupa peluasan sistem dan pemahaman tentang umpan balik selanjutnya. Pada tahap ini tim pengembang menyerahkan sistem kepada CV. Sumber Proteina Bandar Lampung .

3. *Rapid Construction*

Dalam tahapan ini pengujian dilakukan pada tingkatan pengguna (*user*) dari sistem informasi CV. Sumber Proteina, agar dapat diketahui pemahaman terhadap sistem informasi yang dikembangkan.

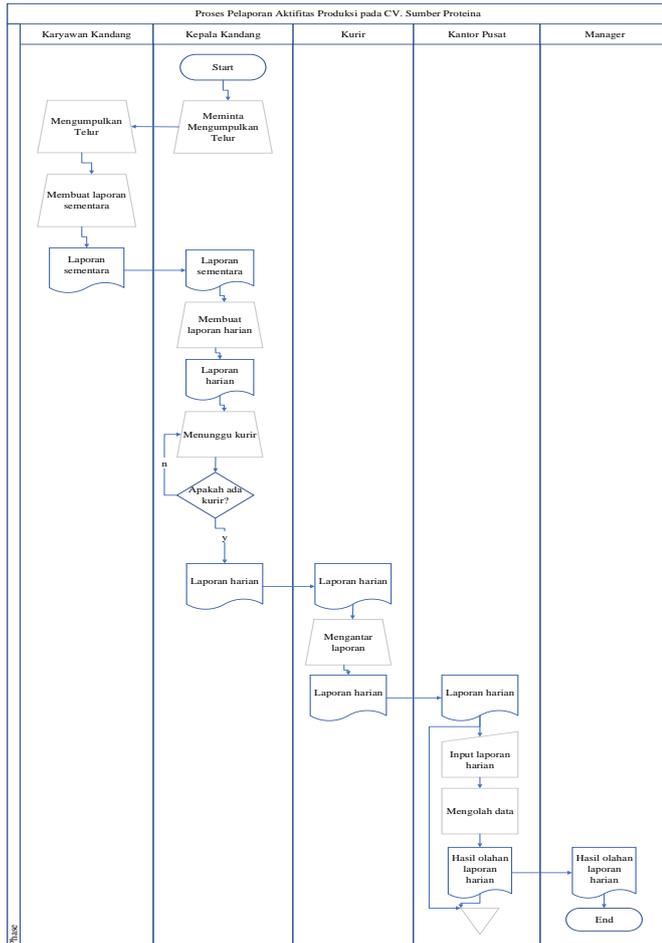
4. *Cutover*

Pada tahapan ini dilakukan pengujian dengan cara menguji setiap fungsi-fungsi, dimana fungsi-fungsi tersebut akan diuji kesesuaiannya dengan hasil yang diharapkan. Pengujian ini biasanya dilakukan berdasarkan spesifikasi yang dianalisa secara informal dan manual. [7]

III HASIL DAN PEMBAHASAN

Pada tahap analisa sistem ini dilakukan analisis dari sistem yang sudah berjalan pada bagian produksi perusahaan CV. Sumber Proteina Bandar Lampung untuk mengetahui alur proses yang berjalan secara terinci digunakan *Document*

Flowchart seperti yang terlihat pada gambar 1.



Gambar 1. Document Flowchart Sistem Informasi yang berjalan.

Dari hasil Analisa dari proses pelaporan produksi yang berjalan di CV. Sumber Proteina, ditemukan beberapa masalah, yaitu :

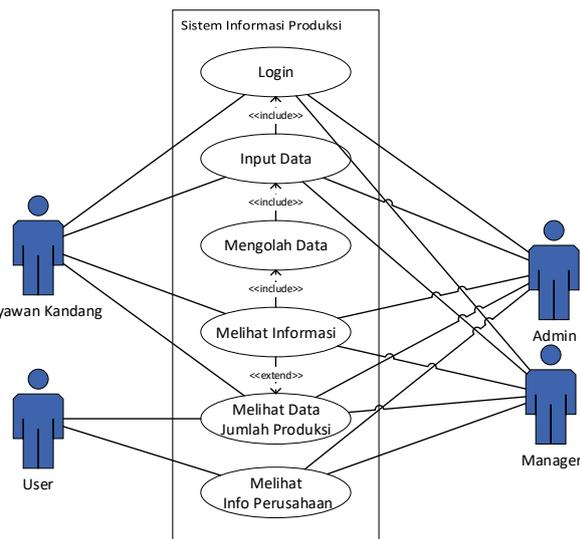
- a. Kepala kandang memerlukan waktu yang cukup lama untuk membuat laporan.

- b. Dalam pengiriman laporan terkendala oleh ada atau tidaknya kurir yang akan mengirim telur.

- c. Laporan produksi telur hanya dapat dilaporkan paling banyak satu kali dalam sehari.

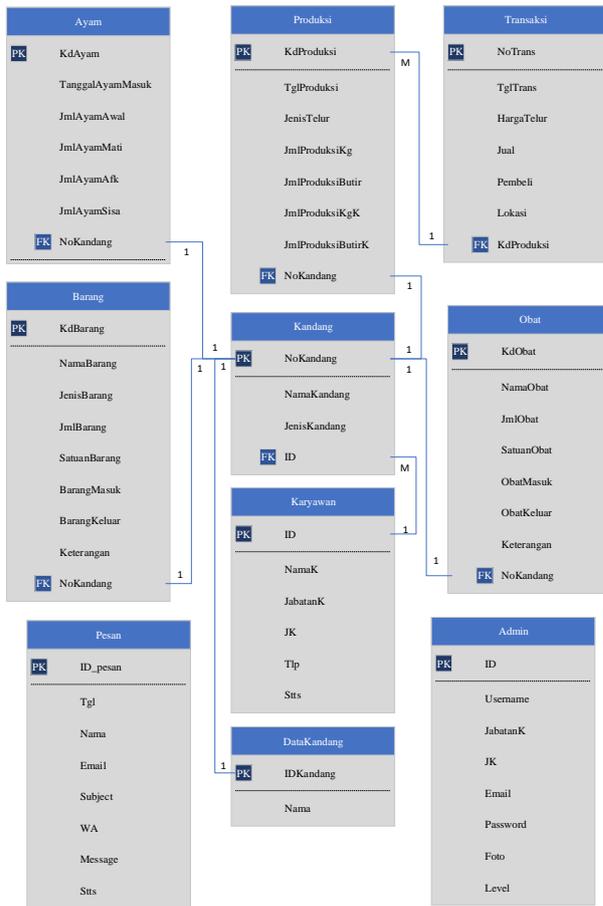
- d. Administrasi memerlukan waktu yang cukup lama untuk mengolah data produksi.

Setelah diketahui permasalahan yang terjadi pada tahap analisis sistem, maka tahap selanjutnya adalah menganalisis alternatif tindakan yang biasa dilakukan. Desain sistem yang akan dibuat meliputi: *Use case*, *Class Diagram*, dan *Sequence Diagram*. Gambar 2 memperlihatkan aktor yang terlibat dalam system informasi ini, yaitu: Karyawan kandang, Admin, Manager dan *user (customer)*, serta aksi yang dapat dilakukan oleh masing-masing aktor.



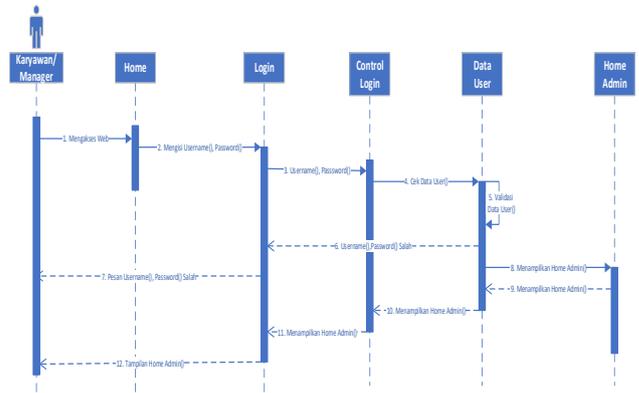
Gambar 2. Use Case yang diusulkan

Adapun *Class diagram* ditunjukkan pada gambar 3, yang menjelaskan kelas-kelas yang terdapat pada sitem informasi ini yaitu sebanyak Sembilan kelas yaitu: Ayam, Produksi, Transaksi, Barang, Kandang, Obat, Karyawan, Pesan, Data Kandang dan Admin.



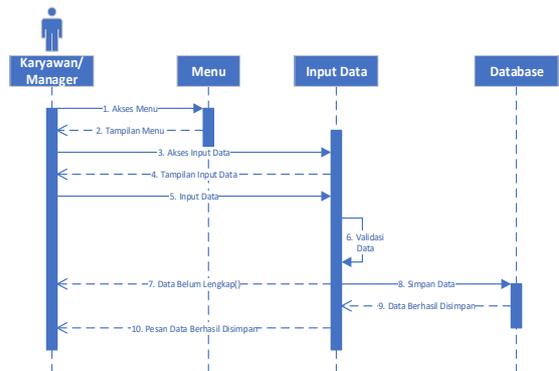
Gambar 3. Class Diagram

Sedangkan *Sequence diagram* untuk proses login ditunjukkan pada gambar 4.



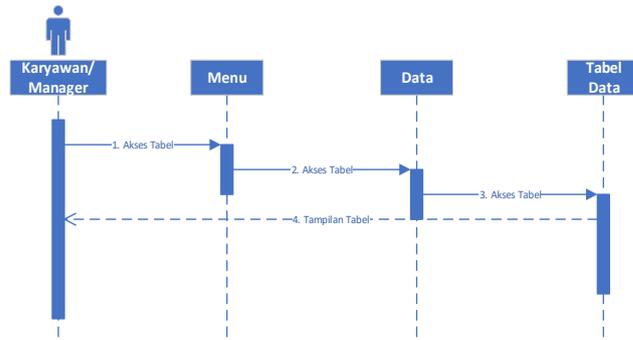
Gambar 4. Sequence Diagram login

Untuk Sequence diagram untuk input ditunjukkan pada gambar 5. Karyawan atau *manager* dapat memasukkan data melalui menu input data, setelah data divalidasi selanjutnya data akan tersimpan pada *database*.



Gambar 5. Sequence Diagram Input

Adapun *Sequence diagram* untuk proses Cek Data ditunjukkan pada gambar 6. Pada proses ini, karyawan atau manager dapat memeriksa data yang terdapat pada table data.



Gambar 6. Sequence Diagram Cek Data

Pengelolaan Data Produksi Berbasis *Web* pada CV. Sumber Proteina Bandar Lampung, memiliki beberapa menu dan halaman informasi, gambar 7 memperlihatkan tampilan awal atau home yang ada pada aplikasi ini.



Gambar 7. Tampilan Home

Untuk halaman info produksi ditunjukkan pada gambar 8, pengguna dapat melihat informasi harga telur, jumlah telur dan jumlah pesan.



Gambar 8. Halaman Info Produksi

Adapun Input Data Harian dilakukan dengan memasukkan data produksi berdasarkan dari *token* yang telah di-*input*-kan sebelumnya yang ditunjukkan pada gambar 9. Pada form input data harian ini, petugas kandang dapat memasukkan data secara *real-time*.



Gambar 9. Menu Input Data Produksi

Adapun Laporan Harian yang menunjukkan tentang Kandang, Jumlah

Hewan Ternak Awal, Jumlah Hewan ternak mati, afkir dan jumlah hewan yang hidup perhari serta jumlah butir telur harian, berat telur setiap kilo per harinya (KG), berat ayam per Gramnya setiap ekor (G/EKOR) dan berat telur setiap gramnya per butir (G/BTR) ditunjukkan pada gambar 10. Laporan yang diproses oleh sistem informasi ini sangat membantu kelancaran operasional dari CV. Sumber Proteina.

KDG	ANAL	MET	AKG	SAK	BATR	GG	WKG	NAMA	JMLAH	G/EKOR	A/BTR	NAMA	JMLAH	FCR
4	2491	1	0	2444	2010	120	81.64	408	330	10432	3478	-	-	0.48
5	2348	1	0	2348	1880	107	77.62	81-1	330	10230	3444	-	-	0.43
6	2821	1	0	2821	2070	126	73.31	81-1	330	10427	3436	-	-	0.41
8	2228	1	0	2227	1880	116	81.81	81-1	330	11216	3428	-	-	0.46
10	2206	1	0	2204	1830	111	82.96	81-1	330	11320	3434	-	-	0.44
11	2810	1	0	2814	2300	115	81.38	81-1	425	10680	3478	-	-	0.48
12	2842	1	0	2844	2490	136	89.84	124-1	425	18338	3482	-	-	0.49
16	2881	0	0	2881	2520	132	87.41	81-1	330	10412	3421	-	-	0.31
18	2874	1	0	2872	2460	136	81.38	81-1	330	14482	3438	-	-	0.42

Gambar 10. Data Laporan Harian Produksi

Pengelolaan Data Produksi telur berbasis web ini memiliki beberapa kelebihan dan keterbatasan. Kelebihan sistem informasi ini, sebagai berikut :

1. Mempermudah karyawan dalam membuat laporan produksi.
2. Menghemat waktu yang diperlukan dalam pengiriman laporan produksi.

3. Terdapat 2 level user yaitu *admin* dan *manager*.
4. Menyebarkan informasi tentang harga telur dan produksi telur yang tersedia pada masyarakat.
5. Masyarakat dapat mengakses halaman *website* terbatas pada halaman awal (*Home*).
6. Masyarakat dapat mengirim pesan kepada *admin* atau *manager* melalui *Menu Contact*.

Sedangkan keterbatasan yang terdapat pada Sistem Informasi ini, sebagai berikut:

1. Belum terdapat penghitung otomatis stok barang dan stok obat.
2. Data produksi yang di-*input*-kan berdasarkan dari nomor kandang atau nama kandang.
3. Jika terdapat kesalahan *input* pada pesan yang dikirim oleh pelanggan, maka, *admin* atau *manager* tidak dapat membalas pesan tersebut.
4. *Notifikasi* pemberitahuan pesan baru yang dikirim oleh pelanggan atau masyarakat hanya dapat dilihat jika mengakses halaman *Home Admin*.

IV. SIMPULAN DAN SARAN

4.1 Simpulan

Berdasarkan pengujian dari system informasi berbasis web ini, dihasilkan beberapa kesimpulan yang merupakan solusi dari kendala yang terdapat pada proses pengolahan dan pelaporan data pada system Informasi Produksi Berbasis Web Pada CV. Sumber Proteina Bandar Lampung. Berikut kesimpulan dari hasil penelitian, antara lain:

1. Sistem ini menghemat waktu dalam pengiriman laporan, karena laporan yang telah di-*input*-kan oleh karyawan di bagian kandang, langsung dapat diakses oleh karyawan kantor dan *manager*.
2. Sistem ini mempermudah karyawan kandang dalam pembuatan laporan, karena data yang di-*input*-kan akan diolah oleh *website*.
3. Sistem ini mempermudah masyarakat untuk mendapatkan informasi mengenai informasi produk di CV. Sumber Proteina, karena informasi produk dapat diakses secara online melalui *website*.
4. Sistem ini mempermudah masyarakat untuk melakukan

pemesanan telur melalui *website*, karena pelanggan atau masyarakat dapat mengirim pesan pada admin melalui layanan *contact*.

4.2 Saran

Berdasarkan hal-hal yang dikemukakan diatas, untuk meningkatkan kualitas Pengelolaan Data Produksi Berbasis Web Pada CV. Sumber Proteina Bandar Lampung yang dibuat maka diberikan saran-saran sebagai berikut :

1. Sistem yang telah dirancang saat ini diharapkan dapat terus di evaluasi, sehingga sistem dapat memberikan pelayanan yang lebih baik lagi.
2. Menambahkan fitur manipulasi data seperti halnya *update* dan *delete* sesuai dengan kebutuhan pengguna.

DAFTAR PUSTAKA

- [1] A. Haidar Mirza, U. Ependi, and F. Panjaitan, "Rekayasa Perangkat Lunak Informasi Kemiskinan," *J. Inform.*, vol. 10, no. 1, pp. 1189–1198, 2016, doi: 10.26555/jifo.v10i1.a3351.
- [2] T. Zaini, "Implementasi Metode

- Basisdata Terdistribusi Untuk Pengelolaan Data Perpustakaan Berbasis Web,” *J. Inform. Darmajaya*, vol. 10, no. 1, pp. 10–23, 2010.
- [3] T. Sutarbi, *Konsep Sistem Informasi*. 2012.
- [4] A. H. Mirza, “Perancangan Basis Data Terdistribusi E-CARGO (Studi Kasus PT. XYZ),” *Semin. Nas. Inform. Yogyakarta*, vol. 2012, no. semnasIF 2012, pp. 1979–2328, 2012.
- [5] A. A. Pradipta, Y. A. Prasetyo, and N. Ambarsari, “Pengembangan Web E-Commerce Bojana Sari Menggunakan Metode Prototype,” *eProceedings Eng.*, vol. 2, no. 1, pp. 1042–1056, 2015.
- [6] Suendri, “Implementasi Diagram UML (Unified Modelling Language) Pada Perancangan Sistem Informasi Remunerasi Dosen Dengan Database Oracle (Studi Kasus: UIN Sumatera Utara Medan),” *J. Ilmu Komput. dan Inform.*, vol. 3, no. 1, pp. 1–9, 2018, [Online]. Available: <http://jurnal.uinsu.ac.id/index.php/algorithm/article/download/3148/1871>.
- [7] W. Wibisono and F. Baskoro, “Pengujian Perangkat Lunak Dengan Menggunakan Model Behaviour Uml,” *JUTI J. Ilm. Teknol. Inf.*, vol. 1, no. 1, p. 43, 2002, doi: 10.12962/j24068535.v1i1.a95.
- [8] Ardiansyah, R. (2019, November). Penggunaan Metode Balance Scorecard Untuk Mengukur Kinerja Pekerjaan Pada PT. Bangun Cipta Karya Pamungkas (PT. BCKP). In *Prosiding Seminar Nasional Darmajaya* (Vol. 1, pp. 78-87).

AUGMENTED REALITY PENGENALAN OBJEK WISATA TAMAN MUMBUL SANGEH BERBASIS ANDROID

Kadek Agus Wirawan¹, I Gede Harsemadi², Ni Kadek Sukerti³

¹²³Prodi Sistem Informasi, Fakultas Informasi dan Komputer,
ITB STIKOM Bali Jalan Raya

Puputan Renon No. 86 Denpasar, Bali, Indonesia tlp. (0361) 244445 fax: (0361) 264773
Aguswira881@gmail.com,harsemadi@stikom-bali.ac.id,dektisamuh@gmail.com

ABSTRACT

Mumbul Park Tourism Object is a tourist attraction located in Sangeh Village, Abiasemal District, Badung Regency. In Mumbul Park there are two temples named Pura Luhur Mumbul (Puru Ulun Siwi) and Pura Ulun Mumbul. In addition to the temple in the Mumbul Sangeh Park Tourism Object, there is a pengelukatan called Pancoran Solas and there is a pond (lake) with thousands of fish. The Mumbul Park Tourism Object has its own charm for tourists, but the lack of information regarding the Mumbul Sangeh Park Tourism Object makes not many people know about the tourist area. Therefore, we need an introduction media regarding the object of the Mumbul Sangeh park which is more modern with Android based Augmented Reality media. This application contains direct information about the history of the Mumbul Park Tourism Object, Pancoran Solas 3D Penglukatan Object, Ulun Danu Temple and the Queen Niang Sakti Statue, Videos and Photos of the introduction of Mumbul Park. The method used in this research is the MDLC (Multimedia Development Life Cycle) method and the C # programming language. The making of this application uses Unity3D, CorelDraw and Adobe Illustrator as the interface designer, Adobe Premiere to edit the introduction video. Based on the results of tests carried out using two testing methods, namely blackbox testing and questionnaires. This application is distributed via Google Drive..

Keywords—Mumbul Sangeh Park, Pancoran Solas, Ulun Danu Temple, Ratu Niang Sakti Statue, Augmented Reality, Android

ABSTRAK

Objek Wisata Taman Mumbul merupakan sebuah objek wisata yang berlokasi di Desa Sangeh, Kecamatan Abiasemal, Kabupaten Badung. Pada Taman Mumbul terdapat dua pura yang bernama Pura Luhur Mumbul (Puru Ulun Siwi) dan Pura Ulun Mumbul. Selain terdapat pura pada Objek Wisata Taman Mumbul Sangeh ini terdapat tempat penglukatan yang bernama Pengelukatan Pancoran Solas dan terdapat sebuah kolam (danau) dengan ribuan ikannya. Objek Wisata Taman Mumbul mempunyai daya tarik tersendiri bagi wisatawan, akan tetapi kurangnya informasi mengenai Objek Wisata Taman Mumbul Sangeh ini membuat tidak banyak orang mengetahui kawasan wisata tersebut. Maka dari itu dalam diperlukan sebuah media pengenalan mengenai objek wisata taman mumbul Sangeh yang lebih modern dengan media Augmented Reality berbasis android.

Dalam aplikasi ini berisikan informasi langsung mengenai sejarah Objek Wisata Taman Mumbul, Objek 3D Penglukatan Pancoran Solas, Pura Ulun Danu dan Patung Ratu Niang Sakti, Video dan Foto pengenalan Taman Mumbul. Metode yang digunakan pada penelitian ini adalah metode MDLC (Multimedia Development Life Cycle) dan Bahasa pemrograman C#. Pembuatan aplikasi ini menggunakan Unity3D, software desain sebagai pembuat desain antar muka software edit video untuk mengedit video pengenalan. Berdasarkan hasil pengujian yang dilakukan menggunakan dua metode pengujian yaitu blackbox testing dan kuesioner. Aplikasi ini didistribusikan melalui Google Drive.

Kata Kunci—Taman Mumbul Sangeh, Pancoran Solas, Pura Ulun Danu, Patung Ratu Niang Sakti, Augmented Reality, Android.

I. PENDAHULUAN

Sumbangan sektor pariwisata terhadap pembangunan ekonomi nasional diukur melalui tolak ukur paling penting yaitu mengenai sumbangan pada neraca pembayaran, pendapatan nasional, penciptaan lapangan kerja serta sektor pariwisata lainnya [1]. Terdapat keistimewaan pada setiap tempat wisata begitu juga dengan Objek Wisata Taman Mumbul Sangeh, dengan daya tarik tersendiri bagi wisatawan. Kurangnya informasi mampu menghambat berkembangnya tempat wisata. Dokumentasi dalam bentuk gambar dan video yang terkait pada objek wisata juga tidak banyak ditemukan. Tempatnya cukup mampu menampung pertunjukan tari dan kegiatan upacara keagamaan [2]. Salah satu media sebagai pendukung pariwisata untuk media promosi dengan digunakan Multimedia Augmented Reality. Merupakan bagian teknologi yang memberikan pengalaman antarmuka yang

unik. Prinsipnya adalah menggabungkan antara objek virtual baik dua atau tiga dimensi melalui objek nyata langsung terlihat didepan kamera. Cara kerjanya dengan mengarahkan kamera smartphone pada gambar yang sudah diberi marker. Pada saat marker telah berhasil teridentifikasi, maka aplikasi akan merender objek 2 dimensi maupun 3 dimensi, kemudian menampilkannya pada layar kamera dan objek yang dirender bisa dilihat dari berbagai arah. Gabungan antara buku biasa dengan teknologi Reality Augmented dinamakan Augmented Reality Book. Memiliki dua komponen utama, diantaranya buku dengan marker bertipe Quick Response Code dan peralatan untuk menangkap marker dengan tujuan menampilkan hasilnya. Merupakan media karena berbentuk bahan cetakan yang berisi marker sebagai target untuk menampilkan objek 3D [3]. Variasi dari virtual

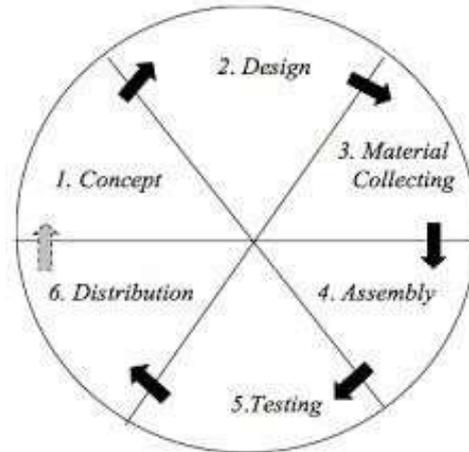
Environment dan lebih dikenal dengan Virtual Reality [4]. Untuk mempercepat link ke website, dial cepat untuk nomor telepon, dan mempercepat pengiriman SMS bisa memanfaatkan Quick Response [5]. Penerapan ke beberapa aplikasi mobile kedalam kedua platform dapat menggunakan SDK Vuforia [6].

Berdasarkan permasalahan yang terdapat pada latar belakang, penulis berharap dapat memberikan informasi tentang pengenalan Objek Wisata Taman Mumbul Sangheh untuk semua masyarakat melalui media Augmented Reality, maka penulis mengusulkan judul proposal tugas akhir dengan judul “Augmented Reality Pengenalan Objek Wisata Taman Mumbul Sangheh Berbasis Android”.

II. METODE PENELITIAN

Metode system dirancang menggunakan metode Multimedia Developmet Life Cycle (MDLC). Terdiri dari 6 tahapan antara lain konsep, desain, material collecting, assembly, testing dan distribution. Pada tahapan konsep terdapat penentuan dari tujuan aplikasi ini. Perancangan dilakukan agar proses spesifikasi perancangan aplikasi dapat dijabarkan secara detail. Dan ini meruakan alur kerja aplikasi Augmented Reality [7]. Penggambaran manual dengan storyboard untuk rangkaian cerita atau deskripsi

setiap scene agar mudah dipahami oleh pengguna, mencantumkan semua objek multimedia dan tautan ke scene lain [8].



Gambar 1. Metode Pengembangan MDLC

Pada tahap penyebarluasan terdaat kegiatan kompresi jika terdapat situasi tidak cukupnya media penyimpanan untuk menampung seluruh aplikasi yang berjalan [9]. Dengan bantuan penyimpanan seperti Google Drive, Dropbox. Class diagram dari aplikasi pengenalan maskot Si Bangun dan Cana berbasis augmented reality. Dimana di dalamnya terdapat penjabaran dari masing-masing proses dari splash screen, main menu, menu Sejarah, Menu Pindai Kamera, menu Galeri dan menu tentang aplikasi. Sangat diperlukan pada tahap pengembangan produk yang telah dibuat sehingga lebih baik melalui proses evaluasi [10].



Gambar 5. Desain Cover Marker Book

Pada bagian deskripsi sejarah ini terdapat informasi mengenai sejarah dari Objek Wisata Taman Mumbul Sangeh.



Gambar 6. Desain Deskripsi Sejarah Objek Wisata Taman Mumbul Sangeh

Penyebarluasan aplikasi maupun marker akan dilakukan melalui melalui link google drive dan link akan langsung terhubung ke penyimpanan google drive yang dimana menyimpan file aplikasi .apk. Aplikasi dan Marker dapat diunduh pada alamat link:

<https://drive.google.com/folderview?id=1ANgTn6GUstR2kZZ9DNBsOI1-N-GdMofl>

IV. SIMPULAN

Kesimpulan dari aplikasi yang telah dibangun, antara lain:

1. Menggunakan media berupa perangkat lunak Unity 3D, Blender 2.8, Adobe

Photoshop, Adobe Premiere, CorelDraw, dan Vuforia, untuk memperkenalkan Objek Wisata Taman Mumbul Sangeh.

2. Terdapat 4 menu yaitu, Sejarah Taman Mumbul, Pindai Kamera, Galeri, dan Tentang. Total objek yang terdapat pada aplikasi augmented reality ini berjumlah 3.

3. Pada hasil pengujian kuesioner didapatkan hasil keseluruhan rata-rata nilai yaitu 3,4 dengan hasil persentase 85% dan dapat disimpulkan bahwa aplikasi ini termasuk kategori sangat setuju.

DAFTAR PUSTAKA

- [1] Ni Putu Ayu Sucahya Dewi, Aplikasi Multimedia Pengenalan Objek Wisata Taman Mumbul Sangeh Berbasis Android. Denpasar; STMIK STIKOM BALI.2020
- [2] Beoang, Dean Dionisius, and Ida Ayu Suryasih. "Identifikasi Potensi Desa Wisata Sangeh, Kabupaten Badung".
- [3] I Nyoman Eva Surya Diana, Media Pengenalan Maskot Si Bangun Dan Cana Pada BAPPEDA Provinsi Bali Berbasis Augmented Reality. Denpasar; ITB STIKOM BALI.2020
- [4] Nugroho, Atmoko, and Basworo Ardi Pramono. "Aplikasi Mobile Augmented Reality Berbasis Vuforia Dan Unity Pada Pengenalan Objek 3d Dengan Studi Kasus Gedung M

- Universitas Semarang." *Jurnal Transformatika* 14.2 (2017): 86-91
- [5] A.A Ngurah Wisnu Widnyana, Aplikasi Augmented Reality Pura Alas Kedaton Berbasis Android. Denpasar; ITB STIKOM BALI. 2020.
- [6] I Gede Putu Surya Gunawan, Aplikasi Pengenalan Barong Landung Pada Museum Bali Berbasis Augmented Reality. Denpasar; ITB STIKOM BALI. 2020.
- [7] Yudi Prasetya, Augmented Reality Pengenalan Ciri-Ciri Pecandu Narkoba Berbasis Platform Android. Denpasar; STIMIK STIKOM BALI. 2020.
- [8] I Wayan Trisna Gunawan, Pengenalan Permainan Olahraga Bola Voli Untuk Siswa Sekolah Dasar Berbasis Augmented Reality. Denpasar; ITB STIKOM BALI. 2020.
- [9] MZ, MM Komarudin. "Pengujian perangkat lunak metode black-box berbasis equivalence partitions pada aplikasi sistem informasi sekolah." *MIKROTIK: Jurnal Manajemen Informatika* 6.1 (2016).
- [10] Nurajizah, Siti. "Implementasi Multimedia Development Life Cycle Pada Aplikasi Pengenalan Lagu Anak-Anak Berbasis Multimedia." *PROSISKO: Jurnal Pengembangan*
- Riset dan Observasi Sistem Komputer 3.2 (2016).

AUDIT E-LEARNING DENGAN FRAMEWORK COBIT 5.0 DI MASA PANDEMI COVID-19

Rini Nurlistiani¹, Neni Purwati², Supri Yanto³

¹²³Fakultas Ilmu Komputer, Informatics & Business Institute Darmajaya
Jl. 2.A. Pagar Alam No. 93, Bandar Lampung - Indonesia 35142
Telp. (0721) 787214 Fax. (0721) 700261
e-mail : rininurlistiani@darmajaya.ac.id, nenipurwati87@darmajaya.ac.id,
suprianto@darmajaya.ac.id

ABSTRACT

The first Coronavirus Disease-2019 (COVID-19) outbreak appeared in Wuhan, China on December 01, 2019 and was declared a pandemic by WHO (world health organization) on March 11, 2020[1]. This condition requires all levels of society to stay at home (WFH), worship, work and study all at home. So that educational institutions are also required to follow government regulations and innovate learning processes that must continue when natural disasters or global pandemics occur through online learning with the aim of improving the quality of learning. Darmajaya Institute of Informatics and Business uses information technology in the form of online learning media called E-Learning for students and lecturers. This of course makes students and lecturers have to adapt in using the information technology. This E-Learning technology is used intensely by students and lecturers (both permanent and external lecturers) of IIB Darmajaya, so that in conditions in the field there are many shortcomings in the E-Learning service at IIB Darmajaya. One way to measure the governance and capability level of E-Learning services at IIB Darmajaya is to conduct a governance audit on the E-Learning services. In this study, a governance framework is needed in the form of Control Objective for Information and Related Technology 5 (COBIT 5.0) which can provide benefits for an agency in achieving strategic goals and optimizing services from information technology.

Keywords—Audit Tata Kelola, E-Learning, COBIT 5.0, Maturity Level, Capability Level.

ABSTRAK

Wabah *Coronavirus Disease-2019* (COVID-19) pertama muncul di Wuhan, China pada tanggal 01 Desember 2019 dan ditetapkan menjadi sebuah pandemic oleh WHO (organisasi kesehatan dunia) pada tanggal 11 Maret 2020. Kondisi ini mengharuskan seluruh lapisan masyarakat agar tetap berada dirumah (WFH), beribadah, bekerja serta belajar semua dilakukan di rumah. Sehingga lembaga pendidikan pun diharuskan mengikuti aturan pemerintah dan melakukan inovasi proses pembelajaran yang harus tetap berjalan ketika terjadi bencana alam atau pandemi global melalui pembelajaran online dengan tujuan meningkatkan mutu pembelajaran. Institut Informatika dan Bisnis Darmajaya menggunakan teknologi informasi berupa media pembelajaran online yang disebut *E-Learning* bagi mahasiswa dan dosen. Hal ini tentu saja membuat mahasiswa dan dosen harus beradaptasi dalam menggunakan teknologi informasi tersebut. Teknologi E-

Learning ini digunakan secara intens oleh mahasiswa dan dosen (baik dosen tetap maupun dosen luar) IIB Darmajaya, sehingga pada kondisi di lapangan terdapat banyak kekurangan yang ada pada layanan E-Learning di IIB Darmajaya. Salah satu cara untuk mengukur tata kelola dan tingkat kapabilitas layanan E-Learning di IIB Darmajaya adalah dengan melakukan audit tata kelola pada layanan E-Learning tersebut. Pada penelitian ini diperlukan suatu *framework* tata kelola yang berupa *Control Objective for Information and Related Technology 5* (COBIT 5.0) yang dapat memberikan manfaat bagi sebuah instansi dalam mencapai tujuan strategis dan mengoptimalkan layanan dari teknologi informasi.

Kata kunci—Audit Tata Kelola, *E-Learning*, *COBIT 5.0*, *Maturity Level*, *Capability Level*.

I. PENDAHULUAN

Wabah *Coronavirus Disease-2019* (COVID-19) pertama muncul di Wuhan, China pada tanggal 01 Desember 2019 dan ditetapkan sebagai pandemic oleh WHO (organisasi kesehatan dunia) pada tanggal 11 Maret 2020[1]. Untuk mencegah penyebaran virus COVID-19 ini, pemerintah telah menghimbau dan mengeluarkan berbagai instruksi dan kebijakan, seperti isolasi mandiri, social distancing hingga pembatasan sosial berskala besar (PSBB) [3]. Kondisi ini mewajibkan seluruh lapisan masyarakat untuk menerapkan kebijakan belajar dan bekerja dari rumah (Work From Home) sejak pertengahan Maret 2020. Sehingga instansi/lembaga pendidikan harus mengikuti aturan pemerintah dengan melakukan berbagai inovasi untuk proses pembelajaran agar tetap berjalan ketika terjadi pandemi COVID-19 melalui pembelajaran online dengan tujuan

meningkatkan mutu pembelajaran[2]. Namun, segala sesuatu dalam harus dapat dimengerti bahwa dalam proses pembelajaran online tersebut tidak akan lepas dari berbagai permasalahan baik internal maupun eksternal yang menjadi hambatan di dalam proses pembelajaran, baik proses pembelajaran online bagi para mahasiswa [4].

Dalam pelaksanaan pembelajaran secara online, Institut Informatika dan Bisnis Darmajaya menggunakan teknologi informasi yang berupa media pembelajaran online yang disebut *Learning Management System* (E-Learning) untuk mahasiswa dan dosen. Hal ini membuat mahasiswa dan dosen harus beradaptasi dengan teknologi tersebut. Sebelumnya E-Learning sudah digunakan sebelum masa pandemi, namun dengan intensitas yang terbatas hanya pada dosen tetap saja dan waktu pelaksanaan adalah 40% dari perkuliahan tatap muka. Namun dengan masa pandemi

COVID-19 yang belum berakhir, teknologi E-Learning ini digunakan secara intens oleh mahasiswa dan dosen (baik dosen tetap maupun dosen luar biasa), serta unit terkait pembelajaran yaitu BAAK, sehingga diperlukan suatu audit tata kelola dalam meningkatkan layanan E-Learning di IIB Darmajaya sesuai dengan kondisi di lapangan.

Salah satu cara mengukur tingkat kapabilitas layanan E-Learning di IIB Darmajaya adalah menggunakan *framework Control Objective for Information and Related Technology 5* (COBIT 5.0) yang dapat memberikan manfaat bagi sebuah instansi dalam mencapai tujuan strategis dan mengoptimalkan layanan dari teknologi informasi. Dalam hal ini proses tata kelola yang dilakukan adalah menganalisa ketersediaan tenaga personil E-Learning yang ada, gangguan sistem (koneksi internet, database/server, hak akses ke E-Learning) serta keterlibatan stakeholder dalam penggunaan E-Learning Darmajaya selama masa pandemi COVID-19.

1.1 Audit

Audit adalah sebuah proses yang sistematis untuk mendapatkan dan menilai bukti-bukti secara objektif, yang berkaitan dengan tindakan-tindakan dan kejadian-kejadian ekonomi untuk menentukan tingkat kesesuaian dengan kriteria yang

telah diterapkan dan mengkomunikasikan hasilnya kepada pihak-pihak yang berkepentingan [5].

1.2 Tata Kelola

Tata kelola (Governance) adalah turunan dari kata “government”, yang berarti membuat kebijakan (policies) yang sejalan dengan keinginan masyarakat a. Sedangkan pengertian “governance” bagi Teknologi Informasi (IT Governance) adalah penerapan kebijakan TI di dalam organisasi agar penggunaan TI (berikut pengadaan dan pelayanannya) diarahkan sesuai dengan tujuan organisasi [6].

1.3 E-learning

E-learning merupakan segala aktivitas belajar yang menggunakan sebuah teknologi elektronik. Melalui e-learning, pemahaman & pengetahuan mahasiswa tentang sebuah materi tidak tergantung pada guru/dosen tetapi dapat diperoleh juga dari media elektronik [7]. Teknologi elektronik yang banyak digunakan misalnya internet, intranet, tape video atau audio, penyiaran melalui satelit, televisi interaktif serta CD-ROM.

1.4 Control Objective for Information and Related Technology (COBIT 5.0)

COBIT 5 adalah *framework* untuk tata kelola dan manajemen pengelolaan TI. *Framework* ini dapat membantu

menciptakan nilai optimal dari penggunaan TI dengan menyeimbangkan antara manfaat yang ada dengan optimalisasi risiko dan penggunaan sumber daya. COBIT 5 memungkinkan TI yang terlibat untuk dapat diatur dan dikelola secara efektif bagi seluruh organisasi yang berkaitan dengan proses bisnis *end-to-end* secara penuh, serta mempertimbangkan TI sesuai dengan kepentingan *stakeholder* internal dan eksternal[8].

1.5 Maturity Level

Maturity Level adalah sebuah metode untuk mengukur level pengembangan manajemen proses, yang memiliki pengertian yaitu dapat mengukur sejauh mana kapabilitas manajemen tersebut. Seberapa baik pengembangan atau kapabilitas manajemen tergantung dari tercapainya tujuan-tujuan COBIT. Adapun salah satu alat pengukuran kinerja suatu sistem teknologi informasi adalah dengan melakukan tingkat kematangan suatu proses (*maturity level*). Arti dari tingkat kematangan suatu proses teknologi informasi COBIT, yaitu mempunyai tingkat kematangan untuk mengontrol proses-proses TI dengan menggunakan metode penilaian (*scoring*) sehingga organisasi dapat menilai semua proses TI yang dimilikinya[9].

1.6 Capability Level

Capability level digunakan untuk mengukur kematangan *IT enterprise*. *Capability level* diadopsi dari ISO/IEC 15504 sebagai standar proses penilaian. Model ini menyediakan pengukuran kinerja dari semua proses pada area *IT governance* maupun manajemen, serta melakukan peningkatan pada area-area yang telah diidentifikasi. Terdapat 6 *Level* kapabilitas proses yang bisa dicapai termasuk *incomplete process* jika prakteknya tidak tercapai sesuai dengan tujuan[10].

Tingkat kematangan dalam suatu pengelolaan dan pengendalian pada proses teknologi informasi dirujuk pada metode evaluasi organisasi sehingga dapat mengevaluasi *capability level* maupun *maturity level*, yang memiliki nilai dari level 0 sampai level 6. Berikut adalah referensi penelitian terdahulu yang relevan dengan penelitian ini:

1. Peneliti: Cherono, Winnie (University Of Nairobi, Kenya), 2015, Judul: An Evaluation Framework for ICT Management Framework Selection in Kenyan Organisations, Pembahasan: Manajemen TIK, COBIT 5, ITIL V3, Uji Validitas & Reliabilitas (SPSS)[11].
2. Peneliti: Zefanya Wahyu Andean , Rudhy Ho Purabaya , I Wayan Widi

- Pradnyana, 2020, Judul: Pengukuran Tingkat Kapabilitas Teknologi Elearning 4.0 Dengan Menggunakan Framework COBIT 5 Pada Domain EDM, APO, dan DSS (Studi Kasus: Universitas Pembangunan Nasional Veteran Jakarta), Pembahasan: Mengukur Tingkat Kapabilitas, E-Learning 4.0, Framework COBIT 5[12].
3. Peneliti: Aan Muslimin, Adi Sapto Raharjo, Sri Lestari, 2020, Judul: Manajemen Resiko Teknologi Informasi Terkait Pandemi COVID-19 Pada SDN 1 Negara Batin Menggunakan Framework COBIT 5 dan ISO/IEC 31000, Pembahasan: Bagaimana Manajemen Resiko dengan Framework COBIT 5 dan ISO/IEC 31000[13].
4. Peneliti: Imriana Aprillia, 2017, Judul: Analisis Kualitas Layanan Teknologi Informasi Menggunakan Framework Control Objectives For Information And Related Technology (COBIT 5) pada Domain Deliver, Service and Support (DSS) (Studi Kasus: Instalasi Rawat Jalan RSUD Salatiga), Pembahasan: Analisis Kualitas Layanan TI, Capability Level, COBIT 5, Domain DSS[14].
5. Rini Nuristiani, and RZ Abdul Aziz, 2018, Judul: Audit Of Information

System Using COBIT 5.0 And ITIL V3 For Information System Of Academic." *Prosiding International conference on Information Technology and Business (ICITB)*, Pembahasan: Audit sistem informasi, COBIT 5, ITIL V3, uji validitas, uji reliabilitas[15].

II. METODE PENELITIAN

Dalam penelitian ini dilakukan tahapan metode penelitian audit tata kelola[16] sebagai berikut:

1. Perencanaan

- 1) Menentukan objek penelitian
- 2) Perumusan masalah
- 3) Studi Pustaka

2. Pemeriksaan Lapangan

- 1) Kuesioner penilaian lapangan
- 2) Pemetaan kuesioner berdasarkan *framework* COBIT 5 objek penelitian.

3. Pelaporan

- 1) Hasil Analisis Kondisi Proses Berdasarkan *Framework* COBIT 5.0
- 2) Penilaian Level *Capability*
- 3) Pengukuran tingkat kematangan (*Maturity level*) dan Analisis GAP

4. Tindak Lanjut

- 1) Hasil Penilaian Level *Capability* dan GAP

- 2) Penentuan *Key Performance Indicator* dan *Critical Success Factors* dari proses-proses TI Berdasarkan *Framework COBIT*
- 3) Rekomendasi hasil proses

III. HASIL DAN PEMBAHASAN

3.1 Perencanaan (*Planning*)

Pada tahap ini dilakukan perumusan masalah dan tujuan penelitian, menentukan objek penelitian yaitu dosen dan mahasiswa yang menggunakan *E-Learning IIB Darmajaya*, serta melakukan studi pustaka (definisi, kerangka kerja, dsb).

3.2 Pemeriksaan Lapangan (*Fieldwork*)

Proses pendistribusikan kuesioner secara online tentang audit layanan e-learning kepada dosen dan mahasiswa, serta melakukan pemetaan *framework COBIT 5.0*. Besarnya sampel dalam penelitian ini ditentukan dengan menggunakan persamaan Slovin sebagai berikut :

$$n = \frac{N}{1 + N.e^2}$$

Keterangan:

- n = jumlah elemen / anggota sampel
 N = jumlah elemen / anggota populasi
 e = error level (tingkat kesalahan)
 (catatan: umumnya digunakan 1 % atau 0,01, 5% atau 0,05, dan 10 % atau 0,1).

Dalam penelitian ini sampel yang digunakan untuk pengisian kuesioner melalui Google Form adalah mahasiswa dan dosen semester Ganjil TA. 2020/2021 sebanyak :

- 100 mahasiswa (masing-masing 15 orang di tiap jurusan dari total populasi 4.113 mahasiswa aktif)
- 70 dosen (masing-masing 5 orang di tiap jurusan dari total 230 dosen aktif).

Adapun desain kuesioner yang dibuat seperti pada table 1 berikut:

Tabel 1. Desain Pertanyaan Kuesioner

Domain Proses	Pertanyaan
APO-07.1	Sumber daya (operator / staff) LMS Darmajaya cukup memadai
APO-07.2	IT personil (personil utama) di LMS sudah ditunjuk secara tepat
APO-07.3	Apakah kompetensi personil IT dikelola dengan baik ?
DSS-04.6	Apakah pelatihan penggunaan LMS telah dilaksanakan dengan baik ?
DSS-05.2	Apakah koneksi internet Anda cukup baik saat menggunakan LMS?
DSS-05.3	Gangguan di LMS (login, koneksi jaringan, upload file) diperbaiki dengan baik & tepat waktu ?
DSS-05.4	Apakah diberikan batasan hak akses ke seluruh menu di LMS Darmajaya ?
DSS-05.6	Apakah database/server LMS sudah cukup memadai ?
DSS-03.4	Apakah operator/staf memberikan solusi & memastikan kendala tsb tidak terjadi lagi ?
MEA-01.1	Apakah Anda selaku stakeholder ikut terlibat dalam pengelolaan layanan LMS ?
MEA-01.2	Apakah penggunaan LMS telah sesuai dengan SOP yang ada ?

1.3 Pelaporan (*Reporting*)

Sebelum melakukan rekapitulasi penilaian *capability level*, dilakukan uji validitas dan reliabilitas data dari responden.

3.3.1 Uji Validitas dan Reliabilitas Audit Tata Kelola

Pengujian validitas dan reliabilitas dilakukan dengan menggunakan aplikasi SPSS 20.

Data kuisioner yang diperoleh dari responden pengguna layanan e-learning ini telah diuji validitasnya dengan kriteria yang sama dengan pengujian domain APO-07, DSS-05, dan MEA-01.

Hasil uji validitas untuk kondisi saat ini (*performance*) dapat dilihat pada tabel 2 berikut:

Tabel 2 Hasil Uji Validitas Kondisi Saat Ini (*performance*)

Domain	Question	r _{hitung}	r _{tabel}	Conclusion
APO-07	Item1	0,849	0,3061	Valid
	Item 2	0,952	0,3061	Valid
	Item 3	0,823	0,3061	Valid
DSS-05	Item1	0,809	0,3061	Valid
	Item 2	0,928	0,3061	Valid
	Item 3	0,588	0,3061	Valid
	Item 4	0,430	0,3061	Valid
MEA-01	Item 1	0,903	0,3061	Valid
	Item 2	0,866	0,3061	Valid

Berdasarkan hasil uji validitas masing-masing sub domain di kondisi saat ini (*performance*) pada tabel 2 dengan menampilkan seluruh item pertanyaan. Hasil yang diperoleh yaitu nilai r_{hitung} >

r_{tabel} dimana nilai r_{hitung} yaitu 0,3061. Dengan demikian seluruh item pertanyaan yang terdapat pada seluruh domain adalah **Valid**. Sedangkan hasil uji reliabilitas pada kondisi saat ini dapat dilihat pada tabel 3 berikut :

Tabel 3 Uji Reliabilitas (*Performance*)

No	Domain	Cronbach's Alpha	N of Items
1	APO-07	0.817	3
2	DSS-05	0.705	4
3	MEA-01	0.688	2

Dari data diatas, didapat hasil pengujian reliabilitas pada domain COBIT 5 adalah diatas 0,60. Hal ini menunjukkan bahwa kuesioner tersebut reliabel serta dimengerti maksud dan tujuannya oleh para responden.

Untuk hasil uji validitas dan reliabilitas untuk kondisi yang diharapkan (*expected*) dapat dilihat pada tabel 4 berikut :

Tabel 4 Hasil Uji Validitas Kondisi Saat Ini (*expected*)

Domain	Question	r _{hitung}	r _{tabel}	Conclusion
APO-07	Item1	0,851	0,3061	Valid
	Item 2	0,929	0,3061	Valid
	Item 3	0,791	0,3061	Valid
DSS-05	Item1	0,762	0,3061	Valid
	Item 2	0,708	0,3061	Valid
	Item 3	0,627	0,3061	Valid
	Item 4	0,829	0,3061	Valid
MEA-01	Item 1	0,901	0,3061	Valid
	Item 2	0,848	0,3061	Valid

Berdasarkan hasil uji validitas masing-masing sub domain di kondisi yang diharapkan (*expected*) pada tabel 3 dengan menampilkan seluruh item pertanyaan.

Hasil yang diperoleh yaitu nilai r hitung $> r$ tabel dimana nilai r hitung yaitu 0,3061. Dengan demikian seluruh item pertanyaan yang terdapat pada seluruh domain adalah **Valid**. Sedangkan hasil uji reliabilitas pada kondisi saat ini dapat dilihat pada table 5 berikut :

Tabel 5 Uji Reliabilitas (Expected)

DOMAIN	PROSES	RATA-RATA	RATA-RATA PROSES
APO07	APO07.01	3.12	3.26
	APO07.02	3.29	
	APO07.03	3.35	
DSS04	DSS04.06	3.46	3.46
	DSS05.02	3.21	3.20
DSS05	DSS05.03	2.99	
	DSS05.04	3.34	
DSS03	DSS05.06	3.27	3.43
	DSS03.04	3.43	
MEA01	MEA01.01	2.86	3.03
	MEA01.02	3.19	
NILAI CAPABILITY(PERFORMANCE)			3.27

Dari data diatas, didapat hasil pengujian reliabilitas pada domain COBIT 5 adalah diatas 0,60, yang menunjukkan bahwa kuesioner tersebut reliabel serta dimengerti maksud dan tujuannya oleh para responden.

3.3.2 Menentukan Maturity Level

Dalam tahap ini dilakukan perhitungan *maturity level* untuk melihat gambaran tata kelola pada perguruan tinggi. Penentuan *maturity level* untuk kondisi saat ini (*performance*) dilakukan melalui pengisian kuesioner *capability level* yang diberikan kepada responden dan telah

ditentukan, dengan hasil penilaian *capability level* sebagai berikut:

Tabel 6 Capability Level (Performance)

No	Domain	Cronbach's Alpha	N of Items
1	APO-07	0.840	3
2	DSS-05	0.774	4
3	MEA-01	0.750	2

Berdasarkan hasil diatas, diperoleh rata-rata *capability (performance)* sebesar **3,27** yang termasuk dalam level **Established** yaitu proses yang telah ada saat ini diimplementasikan menggunakan proses yang telah didefinisikan dan mampu mencapai hasil proses (*outcome*) yang diinginkan. Selanjutnya dilakukan tahapan untuk mengetahui tata kelola pada perguruan tinggi pada kondisi harapan (*expected.*) Ini dilakukan melalui kuisisioner *capability level* yang diberikan kepada responden melalui google form. Berikut daftar hasil *capability* pengolahan data dari responden pada table 7:

Tabel 7 Capability Level (Expected)

DOMAIN	PROSES	RATA-RATA	RATA-RATA PROSES
APO07	APO07.01	4.37	4.44
	APO07.02	4.46	
	APO07.03	4.48	
DSS04	DSS04.06	4.19	4.19
	DSS05.02	4.20	4.20
DSS05	DSS05.03	4.17	
	DSS05.04	4.24	
DSS03	DSS05.06	4.18	4.16
	DSS03.04	4.16	
MEA01	MEA01.01	4.11	4.13
	MEA01.02	4.15	
NILAI CAPABILITY(EXPECTED)			4.22

Berdasarkan hasil diatas, diperoleh rata-rata *capability (expected)* sebesar **4,22** yang termasuk dalam level **Predictable** yaitu proses yang diharapkan mampu dioperasikan dengan batasan yang ditentukan untuk mencapai outcome yang diinginkan. Hal ini memastikan bahwa kinerja proses telah mendukung pencapaian tujuan proses dan tujuan organisasi.

3.3.3 Analisis Kesenjangan (GAP)

Dari hasil analisa di atas, selanjutnya dilakukan penilaian analisis kesenjangan (GAP). Analisis ini menunjukkan kesenjangan/gap antara tingkat kematangan kondisi saat ini dengan tingkat kematangan kondisi yang diharapkan dari penggunaan e-learning IIB Darmajaya, dapat dilihat pada tabel berikut:

Tabel 7 Perbandingan Tingkat Kesenjangan(GAP)

Domain	Proses	P	E	GAP
APO07	MHR	3.26	4.44	1.18
DSS04	MC	3.46	4.19	0.73
DSS05	MSS	3.20	4.20	1.00
DSS03	MP	3.43	4.16	0.73
MEA01	MEA	3.03	4.13	1.10
Rata - Rata				0.948

Keterangan:

MHR: *Manage Human Resources*

MC: *Manage Continuity*

MSS: *Manage Security Services*

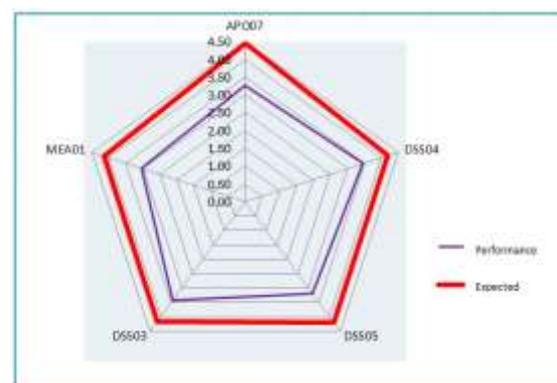
MP: *Manage Problems*

MEA: *Monitor, Evaluate, Assess*

P: *Performance*

E: *Expected*

Berdasarkan tabel diatas, diperoleh rata-rata GAP pada seluruh domain proses yang diteliti sebesar **0,948**. Berdasarkan hasil analisis gap yang diperoleh pada audit tata kelola e-learning Darmajaya dimasa pandemi COVID-19, maka direkomendasikan untuk meningkatkan layanan LMS Darmajaya sekarang dalam mencapai tingkat kemampuan yang diharapkan. Tingkat kemampuan saat ini masih berada pada level 3 yaitu proses yang ada telah diimplementasikan menggunakan proses yang telah didefinisikan mampu mencapai hasil proses (*outcome*) yang diinginkan. Untuk mencapai yang diharapkan di level 4 yaitu proses yang telah diterapkan awalnya dan yang sekarang beroperasi dalam batasan yang telah ditentukan untuk mencapai hasil proses. Adapun analisis GAP juga dapat dilihat pada gambar di bawah ini :



Gambar 1 Grafik Analisis GAP

3.4 Tindak Lanjut (*Follow Up*)

Dalam proses penentuan rekomendasi perbaikan diperlukan pengukuran menggunakan adaptasi dari COBIT 5.0 dan hasil penilaian maturity level serta analisis *GAP* dengan rincian sebagai berikut:

1. Melakukan evaluasi kinerja personil secara terarur untuk memastikan perusahaan memiliki SDM yang memadai
2. Melakukan pelatihan & mengevaluasi secara kontinyu kepada dosen-dosen di IIB Darmajaya tentang penggunaan dan kendala-kendala yang dialami selama penggunaan e-learning.
3. Menentukan hak akses yang tepat kepada dosen, mahasiswa, serta operator layanan terhadap fitur/menu yang ada di e-learning.
4. Mengelola jaringan dengan baik, serta menanggapi resiko/kendala jaringan tepat waktu dan relevan
5. Memperbaiki kendala-kendala yang ada pada e-learning Darmajaya dan memastikan personil IT memiliki rencana yang dikembangkan untuk mencegah insiden terjadi dimasa mendatang.
6. Melibatkan *stakeholder* dalam kendala/masalah yang ada, memonitoring dan mengevaluasi perbaikan yang dilakukan sehingga

meningkatkan efektivitas layanan LMS Darmajaya.

Selanjutnya menentukan *Key Performance Indicator (KPI)* dan *Critical Success Factors* dari proses-proses tata kelola audit sistem informasi layanan e-learning berdasarkan *Framework COBIT 5.0* sebagai berikut:

1. Domain *APO-07 (Manage Human Resource)*
 - a. *Critical Success Factors (CSF)*
 - Mempertahankan staf yang memadai
 - Mengidentifikasi personil-personil IT utama
 - Menjaga keterampilan dan kompetensi personil IT
 - b. *Key Performance Indicators (KPI)*
 - Kompetensi personil IT yang terampil harus mampu mengurangi ketergantungan pada satu individu yang melakukan fungsi pekerjaan penting dalam mencapai tujuan perusahaan
 - Melakukan evaluasi kinerja personil secara teratur untuk memastikan bahwa perusahaan memiliki SDM yang cukup mendukung tujuan dan sasaran perusahaan
 - Mengelola keterampilan dan kompetensi yang dibutuhkan oleh personil IT, memenuhi peran berdasarkan pendidikan, pengalaman yang dimiliki

2. Domain *DSS-04 (Manage Continuity)*
 - a. *Critical Success Factors (CSF)*
 - Melakukan pelatihan
 - Melakukan perencanaan
 - Melakukan review
 - b. *Key Performance Indicators (KPI)*
 - Memberikan semua pihak internal dan eksternal yang berkepentingan pelatihan mengenai prosedur dan peran serta tanggungjawab user jika terjadi gangguan
 - Mengembangkan kompetensi berdasarkan pelatihan yang dilakukan
 - Memantau/mengevaluasi keterampilan dan kompetensi berdasarkan pelatihan yang dilaksanakan
3. Domain *DSS-05 (Manage Security Service)*
 - a. *Critical Success Factors (CSF)*
 - Mengelola jaringan
 - Menjamin keamanan perangkat
 - Mengelola identitas (hak akses) pengguna
 - Mengumpulkan data yang relevan
 - Menanggapi resiko secara tepat
 - Mengelola penanganan resiko
 - Memberikan informasi terkait IT
 - b. *Key Performance Indicators (KPI)*
 - Layanan TI dalam bisnis harus terlindungi demi keamanan pelanggan (*stakeholder*)
- Adanya penentuan hak akses secara jelas dan disepakati berdasarkan kebijakan, terintegrasi dengan kebutuhan bisnis instansi/perusahaan
- Prosedur keamanan yang dilakukan sesuai dan didukung oleh pengelolaan yang baik
4. Domain *DSS-03 (Manage Problems)*
 - a. *Critical Success Factors (CSF)*
 - Mengidentifikasi masalah
 - Menyelediki & mengangkat masalah/kendala
 - Memulai solusi perbaikan yang berkelanjutan
 - b. *Key Performance Indicators (KPI)*
 - Mengidentifikasi dan memulai solusi berkelanjutan yang menjadi akar penyebab masalah dengan baik
 - Memastikan personil IT yang terkena dampak dari gangguan tersebut mengetahui rencana yang akan dikembangkan, untuk mencegah insiden dimasa mendatang.
5. Domain *MEA-01 (Monitor, Evaluate, and Assess)*
 - a. *Critical Success Factors (CSF)*
 - Pendekatan dengan *stakeholder*
 - Mengatur kinerja dan kesesuaian
 - Memproses kinerja dan kesesuaian
 - Menganalisis dan mengevaluasi laporan kinerja

b. *Key Performance Indicators (KPI)*

- Memonitoring dan memvalidasi pengukuran dalam kebijakan bisnis yang sebelumnya melalui SOP yang ada
- Melibatkan *stakeholder* dalam kendala/permasalahan yang ada, sehingga mampu untuk meningkatkan efektivitas layanan e-learning
- Mengevaluasi tindakan perbaikan setiap kali terjadi ketidaksesuaian/kendala penggunaan e-learning
- Melaporkan hasil kinerja layanan e-learning kepada pimpinan / *stakeholder* sebagai evaluasi layanan dimasa yang akan datang.

IV. SIMPULAN

Dari hasil penelitian yang telah dilakukan dapat disimpulkan sebagai berikut:

1. Berdasarkan hasil uji validitas masing-masing sub domain pada kondisi *performance* dan *expected*, diperoleh nilai $r_{hitung} > r_{tabel}$ yang berarti seluruh item pertanyaan (instrumen) yang didapat dari seluruh domain adalah **valid**. Selanjutnya hasil uji reliabilitas menggunakan *Alpha Cronbach* diperoleh koefisien alpha cronbach berkisar dari minimum 0,600 sampai maksimum

0,850 yang menunjukkan bahwa butir pertanyaan tersebut bisa diandalkan atau **reliabel**.

2. Berdasarkan hasil perhitungan *capability level (performance)* diperoleh angka **3,27** yang termasuk dalam level ***Established*** yaitu proses sekarang telah diimplementasikan menggunakan proses yang telah didefinisikan yang mampu mencapai hasil proses (*outcome*) yang diinginkan. Sedangkan *capability level (expected)* diperoleh angka **4,22** yang termasuk dalam level ***Predictable*** yaitu proses yang telah berjalan kemudian dioperasikan dengan batasan yang ditentukan untuk mencapai hasil yang diharapkan. Memastikan bahwa performa proses mendukung pencapaian tujuan proses dan tujuan organisasi.
3. Hasil analisis GAP menunjukkan rata-rata proses sudah baik, dengan selisih GAP antara kondisi saat ini dengan kondisi harapan kurang dari nilai 1 (satu). Artinya semakin kecil analisa kesenjangan pada kondisi saat ini dengan kondisi harapan, maka proses tata kelola TI tersebut semakin baik. Pada proses ini analisis kesenjangan (GAP)

menunjukkan bahwa proses yang telah diterapkan sebelumnya saat ini beroperasi dalam batas-batas yang ditentukan untuk mencapai hasil prosesnya.

Adapun kekurangan dan kelebihan dari penelitian ini adalah:

Kekurangan:

1. Domain yang diteliti tidak mencakup keseluruhan Standar COBIT 5.0.
2. Tidak membahas manajemen resiko teknologi informasi yang digunakan.

Kelebihannya:

Hasil penelitian memberikan rekomendasi perbaikan yang tepat dan akurat karena mengacu pada standar COBIT 5.0 yang telah diakui secara internasional.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada mahasiswa dan dosen yang telah bersedia menjadi responden, dan mengisi kuesioner pada google form, dan kepada pihak pengelola e-learning yang telah memberi dukungan serta atas kesediaannya dijadikan sebagai tempat penelitian ini.

DAFTAR PUSTAKA

[1] WHO Director-General's, *Opening remarks at the media briefing on COVID-19*. 2020.

- [2] A. S. Syarifudin, "IMPELEMENTASI PEMBELAJARAN DARING UNTUK MENINGKATKAN MUTU PENDIDIKAN SEBAGAI DAMPAK DITERAPKANNYA SOCIAL DISTANCING," *J. Pendidik. Bhs. dan Sastra Indones. Met.*, vol. 5, no. 1, 2020, doi: 10.21107/metalingua.v5i1.7072.
- [3] H. S. Siregar, H. Sugilar, and H. Hambali, "Merekonstruksi Alam dalam Kajian Sains dan Agama Studi Kasus pada Masa Pembatasan Sosial Berskala Besar (PSBB) Dampak Covid-19," *Karya Tulis Ilm. UIN Sunan Gunung Djati Bandung*, 2020.
- [4] R. Sanjaya, *Refleksi Pembelajaran Daring di Masa Darurat*. SCU Knowledge Media, 2020.
- [5] R. Tresnawati, "Audit Sistem Informasi Pembayaran Tagihan Rekening Air Bulanan Dengan Menggunakan Kerangka Kerja COBIT 4.1 Pada Perusahaan Daerah Air Minum (PDAM) Kota Bandung," UNIKOM Bandung, 2014.
- [6] K. Surendro, *Implementasi Tata Kelola Teknologi Informasi*. Bandung: Informatika, 2009.
- [7] Rusman, *Pembelajaran Berbasis*

- Teknologi Informasi dan Komunikasi*. Jakarta: Rajawali Press, 2012.
- [8] ITGI, "Cobit 5 : Enabling Process," *United State America*, 2012. .
- [9] ISACA, "COBIT 5," 2012. www.isaca.org.
- [10] ISO/IEC 15504-2, *Software Engineering Process Assessment Part 2: Performing an Assessment*. 2003.
- [11] W. Cheron, "An Evaluation Framework for ICT Management Framework Selection in Kenyan Organisations," 2015.
- [12] I. W. W. P. Zefanya Wahyu Andean , Rudhy Ho Purabaya, "Pengukuran Tingkat Kapabilitas Teknologi Elearning 4.0 Dengan Menggunakan Framework COBIT 5 Pada Domain EDM, APO, dan DSS (Studi Kasus: Universitas Pembangunan Nasional Veteran Jakarta)," 2020.
- [13] S. L. Aan Muslimin , Adi Sapto Raharjo, "Manajemen Resiko Teknologi Informasi Terkait Pandemi COVID-19 Pada SDN 1 Negara Batin Menggunakan Framework COBIT 5 dan ISO/IEC 31000," 2020.
- [14] Imriana Aprillia, "Analisis Kualitas Layanan Teknologi Informasi Menggunakan Framework Control Objectives For Information And Related Technology (COBIT 5) pada Domain Deliver, Service and Support (DSS)," 2017.
- [15] R. N. Listiani and R. Z. A. Aziz, "AUDIT OF INFORMATION SYSTEM USING COBIT 5.0 AND ITIL V3 FOR INFORMATION SYSTEM OF ACADEMIC," ... *Inf. Technol. Bus.*, 2018.
- [16] and F. G. Senft, Sandra, *Information Technology Control and Audit*. CRC Press, 2008.
- [17] Liandi, O., & Fitria, F. (2019). Evaluasi Tata Kelola Framework COBIT 5 pada Dinas Kependudukan dan Pencatatan Sipil. *POSITIF: Jurnal Sistem dan Teknologi Informasi*, 5(2), 111-115.

PEDOMAN PENULISAN JURNAL

1. Naskah yang diusulkan untuk diterbitkan ke dalam jurnal merupakan hasil akhir dari suatu penelitian.
2. Naskah yang diterbitkan adalah naskah yang mempunyai topik ilmu komputer untuk Jurnal Informatika, dengan Ruang Lingkup adalah : Teknologi Informasi, Sistem Informasi, Kecerdasan Buatan, Jaringan Komputer, dan Keamanan Komputer.
3. Naskah yang diterbitkan adalah naskah yang belum pernah diterbitkan atau dipublikasi sebelumnya.
4. Format Penulisan :
 - a. Bahasa : Bahasa Indonesia yang baku menurut Ejaan Indonesia Yang Disempurnakan (EYD)
 - b. Banyak halaman antara 10 – 15 halaman dengan ukuran kertas A4 menggunakan program Microsoft Word dan similariti tidak lebih dari 25%
 - c. Spasi : 1,5 spasi dan ditulis dalam format dua kolom
 - d. Bentuk huruf Time New Roman dengan ukuran huruf 12
 - e. Margin/Batas : batas atas dan kiri 3 cm, batas bawah dan kanan 2,5 cm
 - f. Nama Penulis, dan Afiliansi : ditulis 1 spasi dengan ukuran huruf 10
 - g. Abstrak : Abstract berbahasa Inggris dan Abstrak berbahasa Indonesia ditulis dengan 1 (satu) spasi
5. Sistematika Tulisan : Judul, Nama Penulis, Afiliansi, Abstract dan Abstrak dengan kata kunci maks. 250 kata, 1. Pendahuluan, 2. Metode Penelitian, 3. Hasil dan Pembahasan, 4. Simpulan, Penelitian lanjutan (jika ada/optional), Ucapan Terima Kasih (jika ada/optional), Daftar Pustaka (IEEE Style).
6. Daftar pustaka dibuat secara dengan reference manager mendeley, alfabetis dengan memuat unsur-unsur sebagai berikut : Nama penulis, tahun penerbitan, judul buku/jurnal/proseding, Halaman, Nama penerbit, Kota Asal Penerbit. minimal referensi 10 dan 5 tahun terakhir,
7. Naskah akan diterbitkan pada bulan Juni dan Desember untuk Jurnal Informatika. Naskah yang masuk akan direview oleh tim dan keputusan hasil 2 minggu.
8. Segala sesuatu yang menyangkut perijinan pengutipan atau penggunaan software komputer untuk pembuatan naskah atau hal lain yang terkait dengan HAKI yang dilakukan oleh penulis naskah, berikut konsekuensi hukum yang mungkin timbul karenanya, menjadi tanggung jawab penuh penulis naskah tersebut.
9. Pemeriksaan dan penyuntingan cetak dikerjakan oleh redaksi dan/atau dengan melibatkan penulis. Naskah yang sudah dalam bentuk cetak dapat dibatalkan pemuatannya oleh redaksi jika diketahui bermasalah.
10. Naskah yang masuk ke redaksi akan di seleksi dan hasilnya akan disampaikan kepada penulis dengan kondisi :
 - a. Langsung diterima tanpa perbaikan
 - b. Diterima dengan perbaikan oleh penulis
 - c. Diterima dengan perbaikan oleh redaksi
 - d. Dikembalikan karena kurang memenuhi syarat
11. Isi naskah secara substansi diluar tanggungjawab penerbit dan dewan redaksi/penyunting.
12. Alamat Redaksi : Lembaga Penelitian, Pengembangan Pembelajaran dan Pengabdian pada Masyarakat (LP4M) - Institut Informatika dan Bisnis Darmajaya
Jl. Z.A. Pagar Alam No. 93 A, Bandar Lampung - Indonesia
Telp. 0721-787214 ext: 126
Email jurnal informatika : lp4mjurin@gmail.com

Jurnal Informatika

INSTITUT INFORMATIKA & BISNIS DARMAJAYA
BANDAR LAMPUNG
2021