

**KLASIFIKASI BERITA ONLINE MENGGUNAKAN METODE *ARTIFICIAL
NEURAL NETWORK* (ANN)**

SKRIPSI

Oleh:

UKHTY USWATUN HASANAH

NIM. 18650124



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2022**

**KLASIFIKASI BERITA ONLINE MENGGUNAKAN METODE
*ARTIFICIAL NEURAL NETWORK (ANN)***

SKRIPSI

Oleh:

UKHTY USWATUN HASANAH

NIM. 18650124

Diajukan Kepada:

**Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Malang
untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2022**

HALAMAN PERSETUJUAN

**KLASIFIKASI BERITA ONLINE MENGGUNAKAN METODE
ARTIFICIAL NEURAL NETWORK (ANN)**


SKRIPSI

Oleh:


UKHTY USWATUN HASANAH
NIM. 18650124

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal: 22 November 2022

Pembimbing I



Dr. Cahyo Crysdiyan
NIP. 19740424 200901 1 008

Pembimbing II


Dr. M. Imamudin Lc, M.A.
NIP. 19740602 200901 1 010

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains Dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dr. Fachrud Kurniawan, M.MT., IPM
NIP. 19771020 200912 1 001

HALAMAN PENGESAHAN

**KLASIFIKASI BERITA ONLINE MENGGUNAKAN METODE
ARTIFICIAL NEURAL NETWORK (ANN)**

SKRIPSI

Oleh:

UKHTY USWATUN HASANAH

NIM. 18650124


Telah Dipertahankan di Depan Dewan Penguji
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 8 Desember 2022

Susunan Dewan Penguji

Ketua Penguji	: <u>Dr. Ririen Kusumawati, S.Si, M.Kom</u> NIP. 19720309 200501 2 002	()
Anggota Penguji I	: <u>Dr. Irwan Budi Santoso, M.Kom</u> NIP. 19770103 201101 1 004	()
Anggota Penguji II	: <u>Dr. Cahyo Crysdian</u> NIP. 19740424 200901 1 008	()
Anggota Penguji III	: <u>Dr. M. Imamudin Lc, M.A</u> NIP. 19740602 200901 1 010	()

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains Dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dr. Fachrul Kurniawan, M.MT., IPM
NIP. 19771020 200912 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Ukhty Uswatun Hasanah
NIM : 18650124
Fakultas : Sains dan Teknologi
Jurusan : Teknik Informatika
Judul Skripsi : Klasifikasi Berita Online Menggunakan Metode *Artificial Neural Network* (ANN)

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-banar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiransaya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan Skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 22 November 2022
Yang membuat pernyataan,



Ukhty Uswatun Hasanah
NIM. 18650124

HALAMAN MOTTO

لَئِنْ شَكَرْتُمْ لَأَزِيدَنَّكُمْ

“Sesungguhnya jika kamu bersyukur, pasti Kami akan menambah (nikmat) kepadamu”

(QS. Ibrahim 14:7)

HALAMAN PERSEMBAHAN

الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ

Alhamdulillah, terimakasih sudah sampai sejauh ini.

HALAMAN PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Assalamu'alaikum Wr. Wb.

Puji syukur penulis panjatkan atas kehadiran Allah SWT yang telah melimpahkan rahmat dan kasih sayang-Nya, sehingga penulis mampu menyelesaikan skripsi ini dengan baik dan tepat waktu. Sholawat serta salam semoga tercurahkan kepada Nabi besar Muhammad SAW, atas syafaatnya yang telah menuntun umat manusia dari jalan kebatilan ke jalan yang benar. Semoga kita semua masuk ke dalam golongan yang dituntun Allah SWT serta mendapatkan pertolongan Nabi Muhammad SAW. *Aamiin.*

Selanjutnya penulis mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu dan mendukung terselesaikannya skripsi ini. Ucapan terima kasih ini penulis sampaikan kepada :

1. Prof. Dr. M. Zainuddin, MA selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Fachrul Kurniawan M.MT., IPM, selaku ketua Jurusan Teknik Informatika.
3. Dr. Cahyo Crysdiyan dan Dr. M. Imamudin Lc, MA, selaku pembimbing I dan II yang selalu membimbing penyusunan skripsi ini hingga selesai.
4. Segenap civitas akademika Jurusan Teknik Informatika, terutama seluruh dosen, terima kasih atas ilmu dan bimbingan yang telah diberikan.

5. Bapak Joko Purnomo dan ibu Sri Anah yang telah memberikan do'a dan dukungan baik secara moril dan materi, serta Mufaridjal Humum dan Laila Aisyah Nurhayati yang selalu memberikan semangat sehingga penulis dapat menyelesaikan skripsi.
6. Saudara Teknik Informatika "UFO" angkatan 2018 dan seluruh keluarga besar Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang selalu memberikan semangat dan bantuan selama menyusun skripsi.
7. Semua pihak yang telah membantu dalam penyusunan skripsi ini yang tidak bisa di sebutkan satu persatu.
8. Penulis sendiri karena telah percaya mampu menyelesaikan skripsi ini tanpa harus membandingkan dengan pencapaian orang lain.

Penulis menyadari dalam penyusunan skripsi ini masih banyak terdapat kekurangan. Penulis berharap skripsi ini dapat memberikan manfaat kepada para pembaca terutama terhadap penulis sendiri.

Wassalamu'alaikum Wr. Wb.

Malang, 22 November 2022

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TULISAN	v
HALAMAN MOTTO	vi
HALAMAN PERSEMBAHAN	vii
HALAMAN PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
ABSTRAK	xiv
ABSTRACT	xv
المخلص	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Pernyataan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian	4
BAB II STUDI PUSTAKA	6
2.1 NLP (Natural Language Processing)	9
2.2 Klasifikasi Berita.....	6
2.3 Artificial Neural Network (ANN).....	11
BAB III DESAIN PENELITIAN.....	15
3.1 Koleksi Data.....	15
3.2 Desain Sistem.....	15
3.2.1 Pre processing	16
3.2.2 Term Frequency	20
3.2.3 BufferTerm	22
3.2.4 Artificial Neural Network.....	23
BAB IV UJI COBA DAN PEMBAHASAN.....	33

4.1 Skenario Pengujian.....	33
4.2 Uji Coba	35
4.3 Pembahasan.....	40
BAB V PENUTUP.....	44
5.1 Kesimpulan	44
5.2 Saran.....	44
DAFTAR PUSTAKA	
LAMPIRAN	

DAFTAR GAMBAR

Gambar 3.1 Diagram Blok Sistem	16
Gambar 3.2 Flowchart Pre-processing	16
Gambar 3.3 Flowchart Case folding	18
Gambar 3.4 Flowchart Stopword Removal.....	20
Gambar 3.5 Flowchart Term Frequency	21
Gambar 3.6 Contoh Term Frequency.....	22
Gambar 3.7 Hasil Perhitungan <i>Term Frequency</i>	22
Gambar 3.8 Flowchart Buffer Term.....	23
Gambar 3.9 <i>Architecture Neural Network Multi Layer Perceptron</i>	24
Gambar 3.10 Flowchart Proses Training ANN.....	25
Gambar 3.11 Source Code-Proses ANN.....	26
Gambar 3.12 Source Code- Proses ANN.....	27
Gambar 3.13 Source Code- Proses ANN.....	28
Gambar 3.14 Flowchart Proses Testing ANN	28
Gambar 3.15 <i>Source Code</i> -Fungsi Prediksi	29
Gambar 3.16 Grafik RMSE pada tahap training.....	31
Gambar 3.17 Grafik Accuracy pada tahap training	32
Gambar 4.1 Gambar Perbandingan	40

DAFTAR TABEL

Tabel 3.1 Contoh tokenization	17
Tabel 3.2 Contoh dari Case Folding	18
Tabel 3.3 Contoh Stemming	19
Tabel 3.4 Contoh Stopword	19
Tabel 3.5 Data Training	30
Tabel 4.1 Pembagian Dataset	33
Tabel 4.2 Hasil Klasifikasi Pengujian 90:10.....	36
Tabel 4.3 Confusion Matrix untuk Pengujian 90:10.....	36
Tabel 4.4 Hasil Klasifikasi Pengujian 80:20.....	37
Tabel 4.5 Confusion Matrix untuk Pengujian 80:20.....	37
Tabel 4.6 Hasil Klasifikasi Pengujian 70:30.....	38
Tabel 4.7 Confusion Matrix untuk Pengujian 70:30.....	39

ABSTRAK

Hasanah, Ukhty Uswatun. 2022. **Klasifikasi Berita Online Menggunakan Metode *Artificial Neural Network* (ANN)**. Skripsi. Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (1) Dr. Cahyo Crysdiandian (2) Dr. M. Imamudin Lc, M.A.

Kata Kunci: Klasifikasi Berita, *Artificial Neural Network*, *Natural Language Processing*

Pengklasifikasian dokumen berita bisa dikerjakan dengan cara otomatis menggunakan sebuah sistem, dimana sistem itu bisa mengerjakan sebuah pengklasifikasian artikel berita menurut pola yang sudah diberikan, sistem seperti ini sering disebut dengan istilah *machine learning*. Metode yang diterapkan dalam pembuatan sistem pada penelitian ini merupakan metode *Artificial Neural Network*. Tujuan dilakukannya penelitian ini untuk mengukur nilai *accuracy*, *precision*, *recall*, serta *f-measure*. Menggunakan metode ini karena kemampuannya yang dapat menyelesaikan data kompleks atau kurang tepat, metode ini mampu mendapatkan pola atau *trend* yang sangat kompleks agar dapat dipahami oleh sistem. Kemudian data yang digunakan untuk *input* pada *Neural Network* merupakan data yang telah melalui tahap *preprocessing*. Pada proses *preprocessing* melakukan pendekatan menggunakan *Natural Language Processing* (NLP) yakni *tokenization*, *case folding*, *stemming*, serta *stopword removal*. Selanjutnya data yang telah melalui tahap *preprocessing* akan dihitung frekuensi kemunculan kata, frekuensi inilah yang akan menjadi *input* pada *Neural Network*. Pada pengujian pertama dengan perbandingan 90:10 yaitu *accuracy* 96,29%, *precision* 83,33%, *recall* 83,33%, serta *f-measure* 83,33%. Hasil dari pengujian kedua dengan perbandingan 80:20 yaitu *accuracy* 92,59%, *precision* 66,66%, *recall* 66,66%, serta *f-measure* 66,66%. Hasil dari pengujian ketiga dengan perbandingan 70:30 yaitu *accuracy* 88,27%, *precision* 47,22%, *recall* 47,22%, serta *f-measure* 47,22%. Dari hasil pengujian tersebut data yang digunakan pada saat proses *training* sangat berpengaruh terhadap hasil akhir, sehingga semakin banyak data yang digunakan pada data *training* maka semakin baik pula hasil yang didapatkan. Hal ini disebabkan adanya pola pembelajaran pada sistem.

ABSTRACT

Hasanah, Ukhty Uswatun. 2022. **Online News Classification Using Artificial Neural Network Method**. Thesis. Department of Informatics Engineering, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University Malang. Supervisor : (1) Dr. Cahyo Crysdiان (2) Dr. M. Imamudin Lc, M.A.

Keywords : News Classification, Artificial Neural Network, Natural Language Processing

Classification of news documents can be done automatically using a system, where the system can classify news articles according to a given pattern, a system like this is often referred to as machine learning. The method applied in making the system in this study is the Artificial Neural Network method. The purpose of this research is to measure accuracy, precision, recall, and f-measure values. Using this method because of its ability to resolve complex or inaccurate data, this method is able to obtain very complex patterns or trends so that the system can understand them. Then the data used for input on the Neural Network is data that has gone through the preprocessing stage. In the preprocessing process, the approach is using Natural Language Processing (NLP), namely tokenization, case folding, stemming, and stopword removal. Furthermore, the data that has gone through the preprocessing stage will be calculated for the frequency of word occurrences, this frequency will be the input to the Neural Network. In the first test with a ratio of 90:10, namely 96.29% accuracy, 83.33% precision, 83.33% recall, and 83.33% f-measure. The results of the second test with a ratio of 80:20 are 92.59% accuracy, 66.66% precision, 66.66% recall, and 66.66% f-measure. The results of the third test with a ratio of 70:30 are 88.27% accuracy, 47.22% precision, 47.22% recall, and 47.22% f-measure. From the test results, the data used during the training process greatly influences the final results, so that the more data used in the training data, the better the results obtained. This is due to the pattern of learning in the system.

المخلص

حسنة، أختي أسوة (2022). تصنيف الأخبار عبر الإنترنت باستخدام طريقة الشبكة العصبية الاصطناعية (ANN). البحث الجامعي. قسم الهندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم مالانج. المشرف: (1) الدكتور جاهيو كريسيديان (2) الدكتور محمد إمام الدين، الليسانس، الماجستير.

الكلمات الرئيسية: تصنيف الأخبار، الشبكة العصبية الاصطناعية، معالجة اللغة الطبيعية

يمكن تصنيف المستندات الإخبارية تلقائياً باستخدام نظام التي يمكنها تصنيف المقالات الإخبارية وفقاً لنمط معين، ويسمى هذا النظام بـ *learning machine*. الطريقة المطبقة في صنع النظام في هذا البحث هي طريقة الشبكة العصبية الاصطناعية. الهدف من هذا البحث هو لقياس قيم *recall*، *precision*، *accuracy* و *f-measure*. تستخدم هذه الطريقة نظراً لقدرتها على حل البيانات المعقدة أو غير الدقيقة، فهذه الطريقة قادرة على الحصول على أنماط أو اتجاهات معقدة للغاية حتى يتمكن النظام من فهمها. البيانات المستخدمة للإدخال إلى الشبكة العصبية هي البيانات التي مرت بمرحلة ما قبل المعالجة. في مرحلة ما قبل المعالجة باستخدام نهج معالجة اللغة الطبيعية (NLP)، وهي *tokenization*، *case folding*، *stemming*، و *stopword removal*. بعد ذلك، ستحسب البيانات التي مرت بمرحلة ما قبل المعالجة تكرار حدوث الكلمات، وسيتم إدخال التردد إلى الشبكة العصبية. في الاختبار الأول، بمقارنة 90:10 حصل على النتائج *accuracy 96.29%*، *precision 83.33%*، *recall 83.33%*، و *f-measure 83.33%*. في الاختبار الثاني، بمقارنة 80:20 حصل على النتائج *accuracy 92.59%*، *precision 66.66%*، *recall 66.66%*، و *f-measure 66.66%*. في الاختبار الثالث، بمقارنة 70:30 حصل على النتائج *accuracy 88.27%*، *precision 47.22%*، *recall 47.22%*، و *f-measure 47.22%*. من نتائج الاختبار، تؤثر البيانات المستخدمة أثناء عملية *training* بشكل كبير على النتائج النهائية، لذلك كلما زادت البيانات المستخدمة في بيانات *training*، كانت النتائج التي تم الحصول عليها أفضل. هذا بسبب نمط التعلم في النظام.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi di bidang ilmu komputer saat ini sangat berkembang, salah satunya dengan semakin banyaknya dokumen berita yang tersedia dalam bentuk artikel di internet yang tidak bisa dikendalikan sehingga dapat meningkat kapan saja dan dimana saja. Dengan banyaknya dokumen berita yang bertambah setiap saat, akan menyulitkan editor untuk mengklasifikasikan berita, jika dilakukan secara manual. Proses klasifikasi berita memungkinkan terjadinya kesalahan. Kemampuan klasifikasi pada dokumen-dokumen berita tersebut merupakan pengaturan berita yang umum. Maka untuk saat ini diperlukan klasifikasi dokumen berita, karena pengelompokan berita secara manual sudah tidak efektif lagi. Saat mengklasifikasikan berita secara otomatis, editor tidak perlu lagi memikirkan kategori mana yang termasuk dalam berita (Wongso et al., 2017).

Pengklasifikasian dokumen berita bisa dikerjakan dengan cara otomatis menggunakan kecerdasan buatan agar dapat membuat sebuah sistem, dimana sistem itu bisa mengerjakan sebuah pengklasifikasian artikel berita menurut pola yang sudah diberikan, sistem seperti ini banyak diketahui dengan istilah *machine learning*. Untuk membangun sebuah *machine learning* terdapat banyak metode *soft computing* yang bisa digunakan, ada salah satu metode yang sering dipakai dalam pembelajaran mesin yaitu *Artificial Neural Network* (ANN) atau lebih sering disebut dengan Jaringan Saraf Tiruan. ANN hampir sama dengan manusia,

yaitu belajar berdasarkan contoh dan pola. ANN merupakan hasil dari konfigurasi untuk sistem tertentu, contohnya pada pengenalan pola atau pengklasifikasian data berdasarkan suatu proses pembelajaran. ANN dikembangkan berdasarkan otak manusia, disebut dengan kata tiruan karena ANN mengimplementasikan program pada komputer yang dapat menyelesaikan sejumlah proses perhitungan pada saat pembelajaran (Atiliani, 2013). Adapun alasan mengapa ANN banyak digunakan selain karena kemampuannya yang bisa menyelesaikan data kompleks atau yang kurang tepat, metode ini pun mampu dalam mendapatkan pola atau *trend* yang sangat kompleks agar dapat diakui oleh seseorang ataupun teknik komputasi lainnya (Indah & Setiawan, 2011). Risdianto (2021) melakukan penelitian mengenai pengklasifikasian undang-undang dengan Jaringan Syaraf Tiruan sebagai metode yang digunakan. Pada proses validasi model sebagai pembanding penulis menerapkan metode *K-Fold Cross Validation*. Hasil yang diperoleh penelitian tersebut ialah nilai *accuracy* 98,19%, *precision* 90,95%, *f-measure* 90,95%, serta *recall* 90,95% pada uji orsinil, sedang hasil diperoleh saat menerapkan metode *K-Fold Cross Validation* menghasilkan nilai nilai *accuracy* 98,27%, *precision* 91,38%, *f-measure* 91,38%, serta *recall* 91,38%.

Dalam Qur'an surah Al-Maidah 5:2 Allah *Subhanahu wa ta'ala* berfirman:

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا تَحْلُوا شَعَائِرَ اللَّهِ وَلَا الشَّهْرَ الْحَرَامَ وَلَا الْهَدْيَ وَلَا الْقَلَائِدَ وَلَا آمِينَ الْبَيْتِ الْحَرَامِ يَتَتَوَّعُونَ فِضًّا
 مِنْ رَبِّهِمْ وَرِضْوَانًا ۖ وَإِذَا حَلَلْتُمْ فَاصْطَادُوا ۗ وَلَا يَجْرِمَنَّكُمْ شَنَا نُ قَوْمٍ أَنْ صَدُّوكُمْ عَنِ الْمَسْجِدِ الْحَرَامِ أَنْ
 تَعْتَدُوا ۗ وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ ۗ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ ۗ وَاتَّقُوا اللَّهَ ۗ إِنَّ اللَّهَ شَدِيدُ الْعِقَابِ

“Wahai orang-orang yang beriman, janganlah kamu melanggar syiar-syiar (kesucian) Allah, jangan (melanggar kehormatan) bulan-bulan haram, jangan

(mengganggu) (hewan-hewan kurban) dan (hewan-hewan kurban yang diberi tanda), dan jangan (pula mengganggu) para pengunjung Baitul Haram sedangkan mereka mencari karunia dan ridho Tuhannya! Apabila kamu telah bertahalul (menyelesaikan ihram), berburulah (jika mau) janganlah sekali-kali kebencian(-mu) kepada suatu kaum, karena mereka menghalang-halangi dari Masjidil Haram, mendorongmu hingga melampaui batas (kepada mereka). Tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong menolong dalam berbuat dosa dan permusuhan. Bertakwalah kepada Allah, sesungguhnya sangat berat siksaan-Nya” (QS. Al-Maidah 5: 2).

Pada tafsir Quraish Shihab (TafsirQ, n.d.) dijelaskan “Janganlah kebencian kalian kepada kaum yang menghalangi kalian pergi ke al-Masjid al-Haram, mendorong kalian untuk memusuhi mereka. Hendaknya kalian, wahai orang-orang Mukmin, saling menolong dalam berbuat baik dan dalam melaksanakan semua bentuk ketaatan dan jangan saling menolong dalam berbuat kemaksiatan dan melanggar ketentuan-ketentuan Allah. Takutlah hukuman dan siksa Allah, karena siksa-Nya amat kejam bagi orang-orang yang menentang-Nya. Ayat ini menunjukkan bahwa al-Qur'an telah terlebih dahulu beberapa ratus tahun menganjurkan konsep kerjasama dalam kebaikan, dibanding semua undang-undang positif yang ada.”

Pada ayat *وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَى* Allah menjelaskan “tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa”. Diharapkan penelitian klasifikasi berita ini bisa menjadi kebajikan untuk membuat masyarakat menjadi lebih mudah dalam memilih jenis berita yang sesuai. Karena berita merupakan informasi penting untuk menambah wawasan keadaan yang terjadi dimasyarakat.

1.2 Pernyataan Masalah

Berikut pernyataan masalah dalam penelitian ini:

1. Berapa nilai *accuracy*, *precision*, *recall*, serta *f-measure* dari metode *Artificial Neural Network* dalam mengerjakan pengklasifikasian berita online?
2. Faktor-faktor apa yang dapat mempengaruhi performa pada klasifikasi berita online menggunakan metode *Artificial Neural Network*?

1.3 Tujuan Penelitian

Tujuan yang ada pada penelitian ini:

1. Mengukur nilai *accuracy*, *precision*, *recall*, dan *f-measure* dari metode *Artificial Neural Network* dalam melakukan pengklasifikasian berita online.
2. Menganalisis faktor-faktor apa yang mempengaruhi performa pada klasifikasi berita online menggunakan metode *Artificial Neural Network*

1.4 Batasan Masalah

Batasan masalah yang ada pada penelitian ini:

1. Artikel yang akan digunakan ialah artikel berbahasa Indonesia yang didapat melalui *website* detik.com, kompas.com, cnnindonesia.com, sindonews.com, serta republik.co.id.
2. Kategori yang digunakan pada penulisan ini adalah kategori berita, yaitu politik, teknologi, travel, hiburan, olahraga, pendidikan, ekonomi, hukum, dan kesehatan.

1.5 Manfaat Penelitian

Manfaat yang didapat pada penelitian ini:

1. Redaktur berita agar dapat melakukan pengklasifikasian berita online dengan mudah.
2. Pustakawan universitas serta masyarakat agar mendapatkan berita yang telah diklasifikasikan dengan benar.

BAB II

STUDI PUSTAKA

2.1 Klasifikasi Berita

Klasifikasi merupakan bahasa Prancis yang artinya *classification*. Kata tersebut merujuk pada satu metode yang digunakan untuk melakukan penyusunan data secara teratur atau sama dengan kaidah yang sudah ditetapkan. Menurut KBBI (Kamus Besar Bahasa Indonesia) klasifikasi merupakan menyusun sistem pada kelompok menurut ketentuan-ketentuan yang telah ditetapkan. Klasifikasi ialah proses mengelompokkan data berdasarkan pada tanda-tanda persamaan maupun perbedaan (Subroto, 2019). Sedang berita ialah laporan tentang fakta yang sementara atau sudah terjadi dan disiarkan melalui media baik berupa media cetak, penyiaran, internet, dan dari mulut satu ke mulut lainnya. Dapat disimpulkan klasifikasi berita ialah pengelompokan berita berdasarkan jenis yang sudah ditetapkan. Tujuan dari pengelompokan berita ialah memberikan label pada sebuah berita sesuai dengan jenis berita.

Hidayat dan Rizqi (2020) melakukan penelitian tentang pengklasifikasian berita dengan menggunakan algoritma *Naïve Bayes Classifier* dan *Enhanced Confix Stripping Stemmer*. Peneliti menggunakan *Enhanced Confix Stripping Stemmer* agar dapat memisahkan kata dasar dari *prefiks* serta imbuhan. Sedang untuk proses klasifikasi penulis menggunakan *Naïve Bayes Classifier*. Dataset yang digunakan berasal dari *website* www.jawapos.com, dokumen sebanyak 600 data berita. Dari data-data itu penulis menggunakannya untuk data *training*, data tersebut dibagi rata dalam beberapa kelas berita yaitu Teknologi, Ekonomi,

Olahraga, dan lainnya. Untuk data yang diuji terdapat 40 data *testing*, hasil akurasi yang didapatkan 90%, 90%, 100%, dan 100%, dan mendapatkan nilai akurasi secara keseluruhan 95%.

Rhomadhona dan Permadi (2019) melakukan penelitian mengenai pengklasifikasian berita kriminal dengan metode NBC (*Naïve Bayes Classifier*) menggunakan pengujian *K-fold Cross Validation*. Ada 5 subkategori berita yang digunakan yaitu pencurian, pembunuhan, pemerkosaan, korupsi, serta narkoba. Pada penelitian tersebut dilakukan agar dapat diketahui kemampuan NBC dalam pengklasifikasian berita menggunakan teknik pengujian *K-fold Cross Validation* dan nilai *K* mulai dari 3 hingga 10. Pada penelitian tersebut peneliti mendapat hasil nilai *precision* 98,53%, nilai *recall* 98,44%, dan nilai *accuracy* 99,38%.

Ariadi dan Fithriasari (2015) melakukan penelitian tentang pengklasifikasian dokumen berita bahasa Indonesia menggunakan metode *Support Vector Machine* dan *Naïve Bayes Classifier* dengan CST. Peneliti menggunakan proses *confix-stripping* untuk menemukan kata dasar yang dibutuhkan pada sebuah berita. Pada metode klasifikasi menggunakan dua metode untuk perbandingan, yang pertama *Naïve Bayes Classifier* (NBC) yang umum digunakan pada teks, yang kedua *Support Vector Machine* (SVM) yang melakukan proses sangat baik di data dengan jangkauan yang luas. Peneliti membagi dua data menjadi data *training* serta *testing* yang berjumlah 70:30 data. Hasil dari klasifikasi saat menggunakan algoritma NBC ialah 82,22%. Sedangkan hasil dari klasifikasi menggunakan algoritma SVM ialah 88,1%.

Herwijayanti *et al.* (2018) mengerjakan penelitian mengenai pengklasifikasian berita online, menerapkan metode *cosine similarity* dan *tf-idf*. Peneliti beranggapan *human error* dapat menjadi masalah apabila berita yang *diposting* tidak sesuai dengan kategori dari berita tersebut. Tidak hanya kesalahan dalam *memposting*, mengelompokkan berita juga penting agar dapat memberi kenyamanan pembaca untuk mencari berita sesuai dengan kategori yang diinginkan. Sebelum melaksanakan pengklasifikasian menggunakan *tf-idf* dan *cosine similarity* penulis melakukan *preprocessing* agar dapat memperkecil *term*, kemudian bisa mempercepat proses perhitungan pembobotan menggunakan *tf-idf* dan mempercepat proses *cosine similarity*. Hasil akurasi dalam pengelompokkan berita sebesar 91,25%.

Hanifa *et al.* (2021) melakukan penelitian mengenai perbandingan metode LSTM dan GRU (RNN) untuk mengklasifikasikan berita *hoax* berbahasa Indonesia. Sebelumnya peneliti telah melakukan pengklasifikasian berita *hoax* memakai model RNN yaitu *long short-term memories* (LSTM), kemudian peneliti mencari parameter model sistem yang tepat sehingga dapat menghasilkan kinerja model yang lebih baik. Selanjutnya peneliti menggunakan perbandingan LSTM dan GRU dan mencari parameter terbaik agar dapat mendapatkan kinerja klasifikasi yang optimal. Nilai parameter *epochs* didapatkan: 15 fungsi *optimizer gradient descent* yaitu *rmsprop*, kemudiain untuk *bacth_size*: 64 agar mendapat kinerja optimal menggunakan metode LSTM. Pada penelitian tersebut penulis mendapatkan hasil akurasi sebesar 73% pada metode LSTM serta 64% pada metode GRU.

2.2 NLP (*Natural Language Processing*)

NLP atau *Natural Language Processing* dikenal dengan sistem pengolahan teks, dimana setiap mesin tidak memahami bahasa natural dari manusia sehingga perlu dibantu dengan adanya *Natural Language Processing* (Munasatya & Novianto, 2020). NLP ialah sistem pendekatan komputasi yang digunakan untuk menganalisis sebuah teks (Crowston *et al.*, 2012). Wangsanegara dan Subaeki, (2015) melakukan penelitian mengenai pengimplementasian NLP dalam mengukur akurasi Ejaan Yang Disempurnakan dalam abstrak skripsi, penulis memakai metode *Fuzzy Logic*. Pada penelitian tersebut para penulis mendapatkan hasil presentase kesesuaian sebesar 70% dari 20 data abstrak skripsi. Pengujian tersebut dikerjakan menggunakan dua cara, yang pertama dilakukannya menggunakan cara manual oleh para pembuat aplikasi kemudian selanjutnya dilakukan oleh aplikasi yang sudah dikerjakan. Dalam pengujian tersebut akan dihitung berapa banyak kesalahan tanda baca serta penulisan yang terdapat dalam setiap abstrak skripsi. Pada pengujian yang dilakukan secara berulang dihasilkan nilai yang konsisten tetapi dengan kondisi yang berbeda. Jadi para penulis menyimpulkan bahwa hasil presentase dari pengujian tersebut ialah 70%.

Prasetyo *et al.* (2021) melakukan penelitian mengenai pengimplementasian NLP pada pembuatan *chatbot* dalam program *information technology*. Pada penelitian tersebut para penulis mengerjakan sebuah aplikasi *chatbot* yang bisa membantu para paenggunanya dalam mendapatkan informasi yang berhubungan dengan program *information technology* yang ada di Universitas Surabaya, jurusan Teknik Informatika. Pada *chatbot* tersebut dibuat

menggunakan bahasa inggris untuk menjalankan programnya, sehingga pertanyaan yang diterima semuanya menggunakan bahasa inggris. *Chatbot* yang dibuat menggunakan pendekatan NLP agar dapat memproses pertanyaan yang diuraikan oleh penggunanya dan agar dapat menemukan kunci dari informasi yang dibutuhkan penggunanya. Kemudian sistem memproses pencarian informasi yang diperlukan oleh pengguna, apabila informasi yang diperlukan tidak tersedia sistem akan menjalankan proses *crawling* agar mendapatkan informasi yang diperlukan pengguna. Dibagian validasi penulis menggunakan metode *user validation* dan *cross validation*. Melalui metode *cross validation* penulis mendapatkan nilai akurasi 83%. Pada *user validation* penulis menggunakan 10 data yang didapat dari pengguna untuk dilakukannya uji coba pada sistem sehingga para penulis mendapatkan nilai akurasi sebesar 76%.

Muktafin *et al.* (Muktafin *et al.*, 2020) melakukan penelitian mengenai analisis sentiment terhadap komentar pengguna yang melakukan pembelian barang di *marketplace* Shopee dengan sebuah pendekatan NLP. Pada penelitian tersebut penulis membuat sistem yang dapat menganalisa banyaknya komentar dengan efektif agar lebih mudah dalam mengerti maksud dari komentar tersebut. Pada penelitian tersebut penulis menggunakan algoritma *K-Nearest Neighbor* (KNN) dan *Term Frequency-Inverse Document Frequency* (TF-IDF) dengan sebuah pendekatan NLP. Pengklasifikasian dengan memakai pendekatan NLP menghasilkan nilai akurasi 77%, presisi 80%, serta *recall* sebesar 74%, sedang sistem klasifikasi tanpa menggunakan NLP menghasilkan akurasi sebesar 70%, presisi 80%, serta *recall* 65%.

Rohman *et al.* (2019) melakukan penelitian mengenai pendeteksian emosi yang ada di media sosial menggunakan sebuah pendekatan leksikon serta NLP. Pada penelitian tersebut penulis membuat sistem pendeteksi emosi pada teks, dimana banyak orang yang condong mengutarakan emosinya melalui media sosial baik menggunakan bentuk video, teks, ataupun gambar. Penulis memakai *EmoLex* sebagai leksikon agar dapat melakukan pendeteksian emosi pengguna yang ada di teks, *EmoLex* memiliki 105 bahasa termasuk bahasa Indonesia dengan memiliki 14.182 kata selanjutnya ditingkatkan menggunakan pencarian sinonim dengan memakai *kataglo API*. Pengujian tersebut menunjukkan pendekatan leksikon dapat mendeteksi 55% atau sekitar 15.357 kata dari 28.696 kata yang sudah didapatkan melalui *update* status para pengguna *facebook*. Kemudian *update* status itu diperbaiki menggunakan pendekatan NLP, hasil perbaikannya dinilai menggunakan leksikon yang sudah dibuat 26 dari 100 *update* status yang dapat dikenal dengan *label* emosinya. Sehingga hasil akurasi label emosi sebesar 61,53%.

2.3 Artificial Neural Network (ANN)

Kanker payudara merupakan jenis kanker yang sering diderita oleh kaum hawa. Sampai saat ini cara yang sering digunakan untuk mengobati kanker payudara ialah dengan melakukan pembedahan dan apabila perlu setelah pembedahan dilakukannya kemoterapi ataupun radiasi. Tetapi pengobatan seperti itu tidak akan memberikan akibat yang cukup baik bagi penderita kanker stadium akhir. Jadi alangkah baiknya jika penyakit tersebut dapat prediksi lebih awal, agar akibat buruk yang kemungkinan akan ditimbulkan oleh penyakit tersebut dapat

ditangkal. Zamani *et al.* (2012) melakukan penelitian mengenai pengimplementasian algoritma genetika pada struktur *Backpropagation Neural Network* pada pengklasifikasian kanker payudara. Di penelitian tersebut para peneliti membandingkan tiga metode untuk mengukur akurasi yang dihasilkan dari ketiga metode tersebut. Tiga metode tersebut diantaranya, NN yang menggunakan parameter algoritma genetika, *Naïve Bayesian*, dan NN dengan *Association Rules*. Penulis melakukan uji coba menggunakan metode *10 fold cross validation* dan memperoleh hasil akurasi dari metode *Neural Network* yang menggunakan parameter algoritma genetika ialah 97,00%, hasil akurasi dari metode *Naïve Bayesian* ialah 96,24%, dan hasil akurasi metode *Neural Network* dengan *Association Rules* ialah 95,6%.

Penyakit jantung merupakan penyakit yang mematikan, penanganan penyakit tersebut harus dilakukan sesegera mungkin karena kemungkinan dapat terjadi kapan saja pada penderitanya. Hal-hal yang dapat memicu penyakit tersebut muncul harus diketahui mulai awal agar resiko terjadinya serangan seketika dapat diatasi ataupun diminimalisir menggunakan beberapa cara misalnya dengan menjalani pola hidup yang sehat serta berolahraga yang teratur. Nawawi *et al.* (2019) melakukan penelitian mengenai komparasi algoritma NN dan *Naïve Bayes* agar dapat mendeteksi penyakit jantung. Pada penelitian tersebut peneliti membandingkan antar dua algoritma untuk mengukur tingkat akurasi terhadap faktor-faktor yang memiliki pengaruh agar dapat mendeteksi penyakit jantung. Hasil akurasi pada penelitian tersebut untuk metode *Neural Network* ialah 84,52%, dan hasil akurasi untuk metode *Naïve Bayes* ialah 79,88%. Dengan

menggunakan 14 variabel yang sama dalam mendeteksi penyakit jantung. Algoritma NN jauh lebih baik dibandingkan dengan algoritma *Naïve Bayes*.

Rini *et al.* (2016) melakukan penelitian mengenai pengklasifikasian dengan metode *Hybrid Bayesian-Neural Network* untuk studi kasus pada identifikasi virus yang ada di komputer. Pada penelitian tersebut peneliti melakukan suatu identifikasi virus dan menggunakan dua metode sekaligus yaitu *Naïve Bayes* dan NN. *Naïve Bayes* digunakan untuk klasifikasi awal dengan membagi dua fitur diantaranya virus dan fitur bukan virus. Sedang untuk metode *Neural Network (perceptron)* digunakan jika sudah dalam jenis virus kemudian akan mengklasifikasikan dalam dua jenis virus yaitu *trojan* atau *worm*. Dari lima kali pengujian data *training* dan *testing* didapatkan hasil *recognition rate* sistem 94,12%.

Andika *et al.* (2019) melakukan penelitian mengenai pengklasifikasian penyakit *Pneumonia* dengan metode *Convolutional Neural Network* dan optimasi *adaptive momentum*. *Pneumonia* merupakan infeksi dari bakteri *Streptococcus pneumonia* yang dapat menyebabkan radang pada kantung udara di paru-paru. *Pneumonia* merupakan penyakit yang dapat menyebar melalui udara, *pneumonia* ialah penyakit berbahaya karena bisa menyebabkan kematian, karena itu perlunya dilakukan deteksi dini menggunakan radiografi dada gambar untuk menentukan gejala yang disebabkan oleh *pneumonia*. Diagnosis menggunakan radiografi dada umumnya masih dikerjakan secara manual oleh para tenaga medis atau dokter, sehingga memerlukan waktu yang lama bahkan sulit untuk mendeteksi penyakit tersebut. *Convolutional Neural Network (CNN)* merupakan metode pembelajaran

yang mengambil kinerja neuron otak manusia untuk mengklasifikasikan citra gambar. Pada penelitian yang dilakukan peneliti berharap CNN dapat membantu mengklasifikasikan *pneumonia* berdasarkan pada gambar radiografi dada. Data yang dipakai pada penelitian tersebut ialah data *Labeled Optical Coherence Tomography* (OCT) dan *Chest X-Ray Image* untuk melakukan pengklasifikasian data sebanyak 5860 gambar yang masuk kedalam dua jenis yaitu normal dan *pneumonia*, peneliti juga menggunakan sampel data sebanyak 2400 orang yang diambil menggunakan *simple random sampling*. Hasil dari klasifikasi yang dibangun ialah 99,98% untuk 100 data *training*, dan 97% untuk data validasi. Data yang sudah dibentuk kemudian di uji menggunakan data *test* dan mendapatkan hasil akurasi sebesar 78% dengan rincian 98 data pada rincian paru-paru normal diklasifikasikan secara benar dan 2 salah, sedang untuk paru-paru yang memiliki kondisi *pneumonia* dihasilkan 42 data salah dan 58 data benar.

BAB III

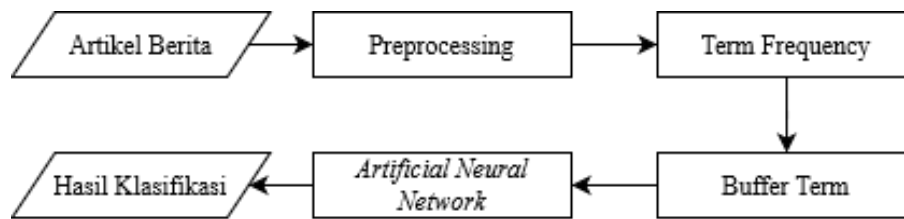
DESAIN PENELITIAN

3.1 Koleksi Data

Ada dua tipe data yang dipakai pada suatu penelitian, yaitu primer serta sekunder. Tipe data primer merupakan data yang diukur oleh peneliti sendiri, sedang tipe sekunder merupakan data yang didapatkan dari sumber lain. Pada penelitian ini data yang akan digunakan merupakan tipe sekunder. Data tersebut diperoleh dari berbagai *website* penyedia artikel berita online yang dapat dilihat oleh siapapun. Sumber data artikel pada penelitian ini didapatkan melalui *website* detik.com, kompas.com, cnnindonesia.com, sindonews.com, serta republika.co.id pada situs-situs tersebut mereka menyediakan berbagai jenis berita yang masuk ke dalam kategori politik, teknologi, travel, hiburan, olahraga, pendidikan, ekonomi, hukum, dan kesehatan. Data yang digunakan berjumlah 360 data artikel berita, dimana masing-masing kategori terdapat 40 artikel berita. Artikel berita yang digunakan diambil dari bulan Juni 2022.

3.2 Desain Sistem

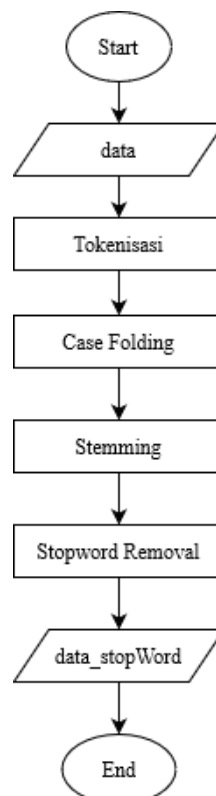
Pada penelitian ini alur desain sistem bisa dilihat pada Gambar 3.1. Melalui beberapa proses yakni artikel berita, *pre-processing*, *term frequency*, *buffer term*, *artificial neural network*, serta hasil klasifikasi.



Gambar 3.1 Diagram Blok Sistem

3.2.1 Pre processing

Ditahap *pre processing*, data artikel berita yang sudah didapatkan dimasukkan dalam repository lalu di ekstrak teksnya melewati empat proses yakni *tokenization*, *case folding*, *stemming*, serta *stopword removal*. Dapat dilihat Gambar 3.2 proses dari *pre processing*. Berikut merupakan penjelasan mengenai *tokenization*, *case folding*, *stemming*, serta *stopword removal*:



Gambar 3.2 Flowchart Pre-processing

1. *Tokenization*

Proses *tokenization* merupakan proses untuk membagi teks kalimat menjadi token-token atau bagian-bagian tertentu. Kata lainnya yaitu, *tokenization* ialah bentuk proses segmentasi pada kalimat. Biasanya acuan yang digunakan sebagai pemisah antar token ialah spasi dan tanda baca. Tabel 3.1 merupakan tabel contoh dari kalimat yang sebelum dan sesudah dilakukannya *tokenization*.

Tabel 3.1 Contoh *tokenization*

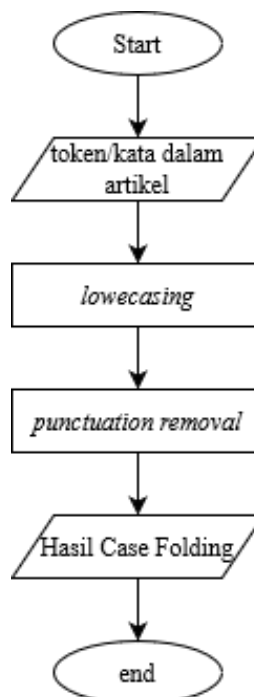
Kalimat yang belum di <i>tokenization</i>	“Pada penelitian ini menggunakan metode Artificial Neural Network dan terdapat sembilan kategori yang digunakan sebagai pengklasifikasian berita”
Kalimat yang sudah di <i>tokenization</i>	“Pada”, “penelitian”, “ini”, “menggunakan”, “metode”, “Artificial”, “Neural”, “Network”, “dan”, “terdapat”, “sembilan”, “kategori”, “yang”, “digunakan”, “sebagai”, “pengklasifikasian”, “berita”

2. *Case Folding*

Case folding ialah proses merubah bentuk pada teks yang ada dikalimat menjadi bentuk yang sama. Pada tahap ini mulai dari merubah huruf-huruf yang ada pada suatu dokumen jadi *lowercasing* (huruf kecil) dikarenakan dengan huruf yang besar dan kecil diduga tidak mempunyai kontras, sehingga bentuk kata yang terdapat pada dokumen diubah jadi *lowercasing*. Pada tahap *case folding* hanya mengubah karakter dari “a” sampai “z”, yang bukan karakter tersebut akan dianggap delimiter dan dihilangkan. Tabel 3.2 merupakan contoh kalimat yang belum dan sudah dilakukannya *case folding*. Gambar 3.3 merupakan proses dari *case folding*.

Tabel 3.2 Contoh dari *Case Folding*

Kalimat yang belum melalui <i>Case folding</i>	“Pada penelitian ini menggunakan metode Artificial Neural Network dan terdapat sembilan kategori yang digunakan sebagai kelas berita”
Kalimat yang sudah melalui <i>Case folding</i>	“pada”, “penelitian”, “ini”, “menggunakan”, “metode”, “artificial”, “neural”, “network”, “dan”, “terdapat”, “sembilan”, “kategori”, “yang”, “digunakan”, “sebagai”, “kelas”, “berita”

Gambar 3.3 *Flowchart Case folding*

3. *Stemming*

Stemming ialah proses agar menemukan kata dasar dari kata baku yang mempunyai imbuhan. Memiliki tujuan yang sama dengan *Case Folding* yaitu agar dapat menghilangkan ambiguitas serta kesalahan dalam proses perhitungan *term frequency* disebabkan perbedaan kapasitas kata. Tabel 3.3 merupakan contoh kata yang telah melewati proses *stemming*.

Tabel 3.3 Contoh *Stemming*

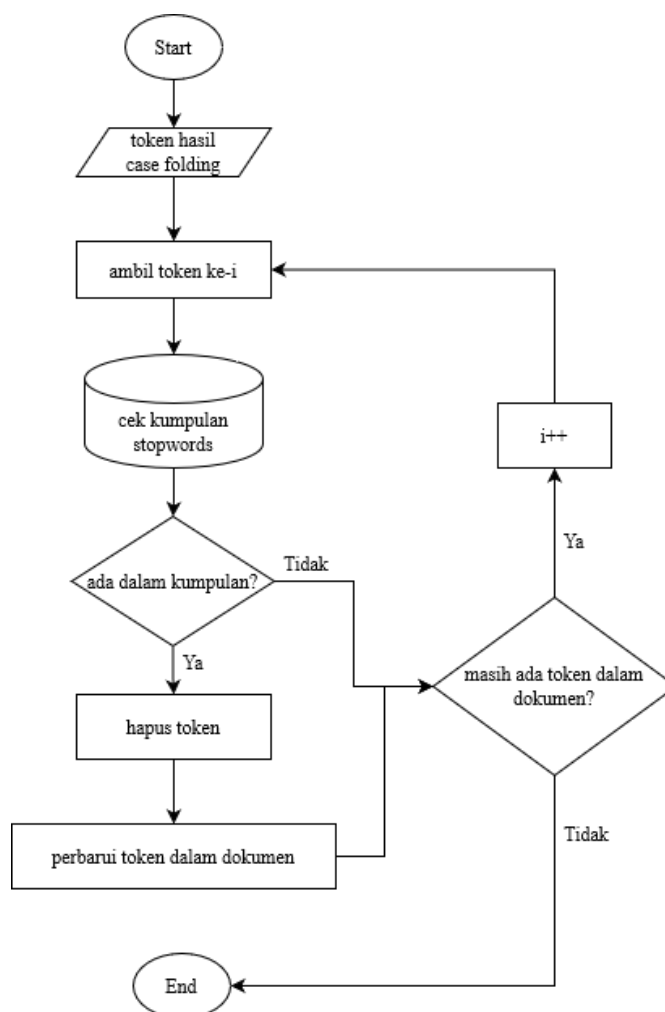
Kata	Imbuhan	Hasil
Sebuah	Se (Awal)	Buah
Menyuci	Me (Awal)	Cuci
Mengubah	Me (Awal)	Ubah
Penyiar	Pe (Awal)	Siar
Mendapatkan	Me (Awal) dan kan (Akhir)	Dapat
Memarahi	Me (Awal) dan i (Akhir)	Marah

4. *Stopword removal*

Stopword removal ialah sebuah proses memfilter atau melepaskan kata-kata yang sering muncul dan dirasa tidak mempunyai keuntungan yang unik dalam kalimat. *Stopword* diartikan juga dengan kata yang sering dijumpai pada suatu kalimat tanpa bergantung pada bahasan tertentu. Tabel 3.4 merupakan contoh kata yang telah melewati proses *stopword removal*. Daftar lengkap dari *stopword* terdapat pada Lampiran I. Pada penelitian ini *stopword* yang digunakan ialah *stopword* dari penelitian Tala (2013). Gambar 3.4 merupakan proses *stopword removal*.

Tabel 3.4 Contoh *Stopword*

ada	ialah	adanya	yang
lah	lain	yang	seandainya
bagaikan	sebagai	bahwa	mau



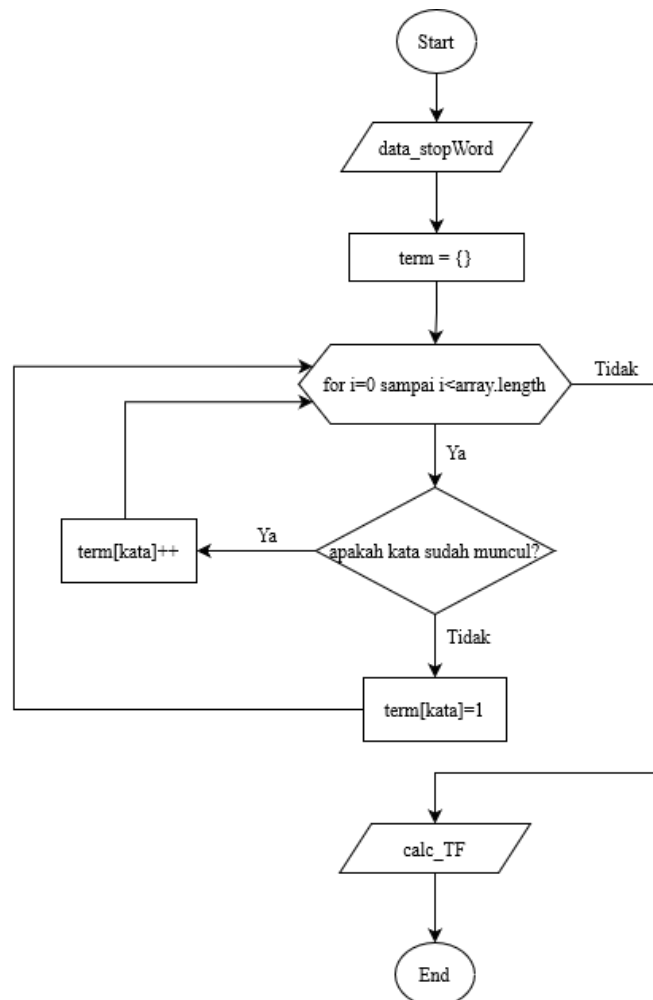
Gambar 3.4 *Flowchart Stopword Removal*

3.2.2 *Term Frequency*

Selanjutnya pada proses pengklasifikasian berita, token akan diganti membentuk vektor menggunakan fitur ekstraksi *Term Frequency* (TF). TF merupakan algoritma berdasarkan nilai statistik yang memperlihatkan kemunculan sebuah kata dalam dokumen. Munculnya suatu kata harus dihitung agar dapat mengetahui kata apa yang mewakili suatu kategori berdasarkan frekuensi kata yang tinggi atau rendah. Sebelum menghitung frekuensi kata, setiap kata dalam dokumen perlu dihitung dulu. Jumlah tiap kata dalam dokumen akan disimpan

didalam objek dengan kata sebagai kunci serta jumlah sebagai nilai objek.

Gambar 3.5 menjelaskan tentang alur dari *term frequency*.



Gambar 3.5 Flowchart Term Frequency

Jika term yang ada pada dokumen telah selesai didapatkan, maka frekuensi term akan dihitung menggunakan rumus Persamaan 3.1.

$$TF \ term_i = \frac{N \ term_i}{\sum_{i=1}^n N} \quad (3.1)$$

Keterangan,

$TF \ term_i$: Frekuensi *term*

$N \ term_i$: Total *term* ke 1, *term* ke 2, ..., *term* ke N

$\sum n$: Jumlah semua *term* yang ada didalam artikel

Untuk contohnya, pada array berisi kata dasar yaitu ['kerja', 'kantor', 'pergi', 'pergi'], setelah dihitung dan disimpan di object maka akan data akan tersimpan seperti yang terdapat pada Gambar 3.6.

```
term = {
  kerja : 1,
  kantor : 1,
  pergi : 2
}
```

Gambar 3.6 Contoh *Term Frequency*

Jika semua *term* sudah selesai dihitung maka selanjutnya *term* akan dihitung frekuensinya, hasilnya seperti yang terdapat pada Gambar 3.7.

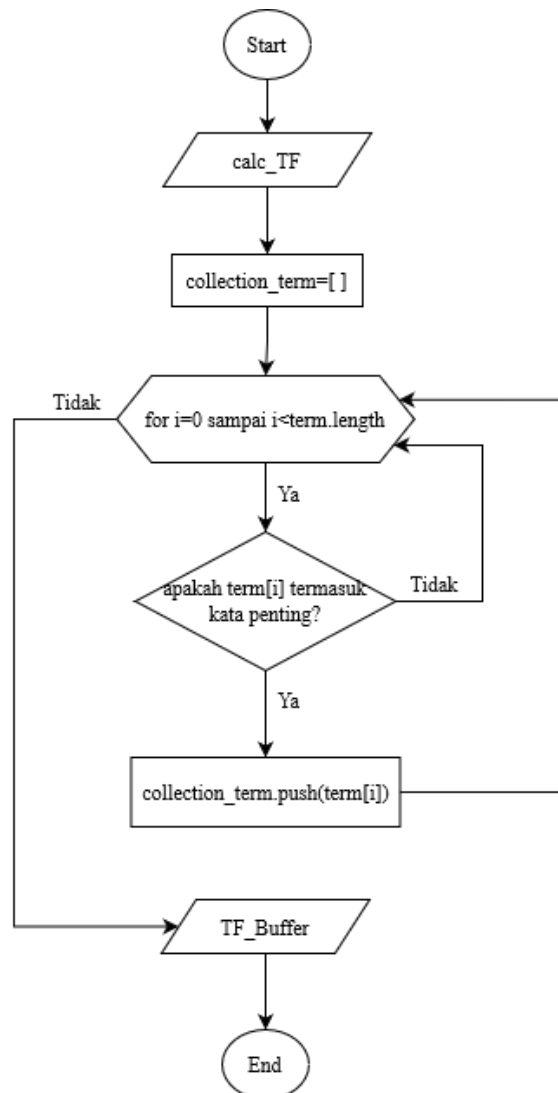
```
TF term.1 = 1/4 = 0.25
TF term.2 = 1/4 = 0.25
TF term.3 = 2/4 = 0.5
```

```
term = {
  kerja : 0.25,
  kantor : 0.25,
  pergi : 0.5
}
```

Gambar 3.7 Hasil Perhitungan *Term Frequency*

3.2.3 *BufferTerm*

Buffer Term ialah data yang sudah melalui tahap *preprocessed*, data tersebut nanti dipakai untuk inputan pada *Neural Network*. *Buffer Term* sendiri disusun dari kata-kata unik dari setiap kategori berita. Gambar 3.8 menjelaskan alur dari pembuatan *buffer term*.

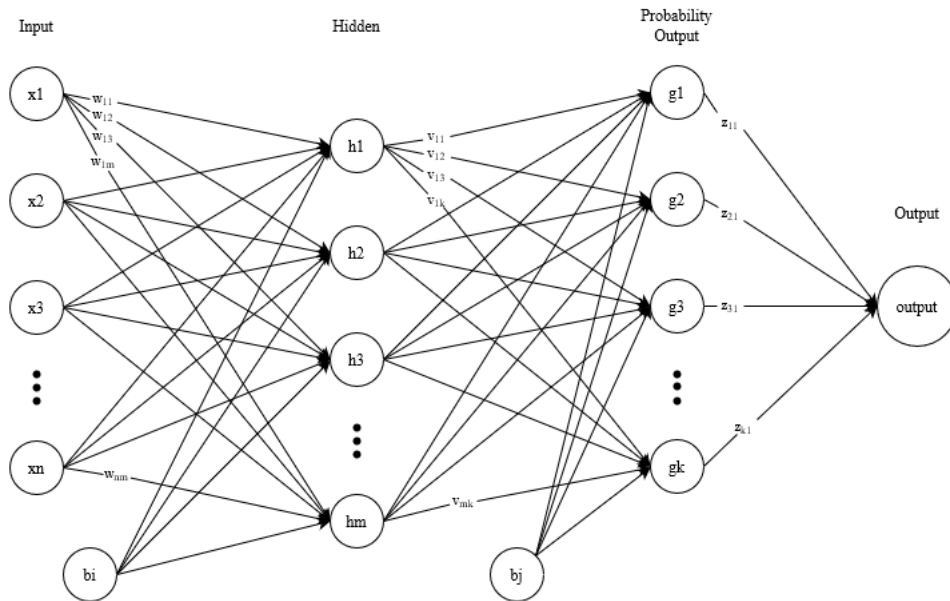


Gambar 3.8 Flowchart Buffer Term

3.2.4 Artificial Neural Network

Dalam penelitian ini, penulis menggunakan *Multi Layer Perceptron*. Arsitektur dari *Neural Network* dapat dilihat pada Gambar 3.9. Arsitektur dari *Artificial Neural Network* pada penelitian ini mempunyai empat layer diantaranya *layer input*, *layer hidden*, *probability output*, terakhir *layer output*. Pada *layer input* disusun dari beberapa node, node tersebut nanti akan diisi dengan frekuensi kemunculan kata atau data yang telah melalui tahap *preprocessed*. Rumus yang

digunakan pada arsitektur *Artificial Neural Network* terdapat pada Persamaan 3.3 hingga Persamaan 3.5, untuk rumus fungsi aktivasi sigmoid terdapat pada Persamaan 3.2.



Gambar 3.9 Architecture Neural Network Multi Layer Perceptron

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (3.2)$$

$$h_j = \text{sigmoid}\left(\sum_{i=1}^n x_i \cdot w_{ij} + b_i\right) \quad (3.3)$$

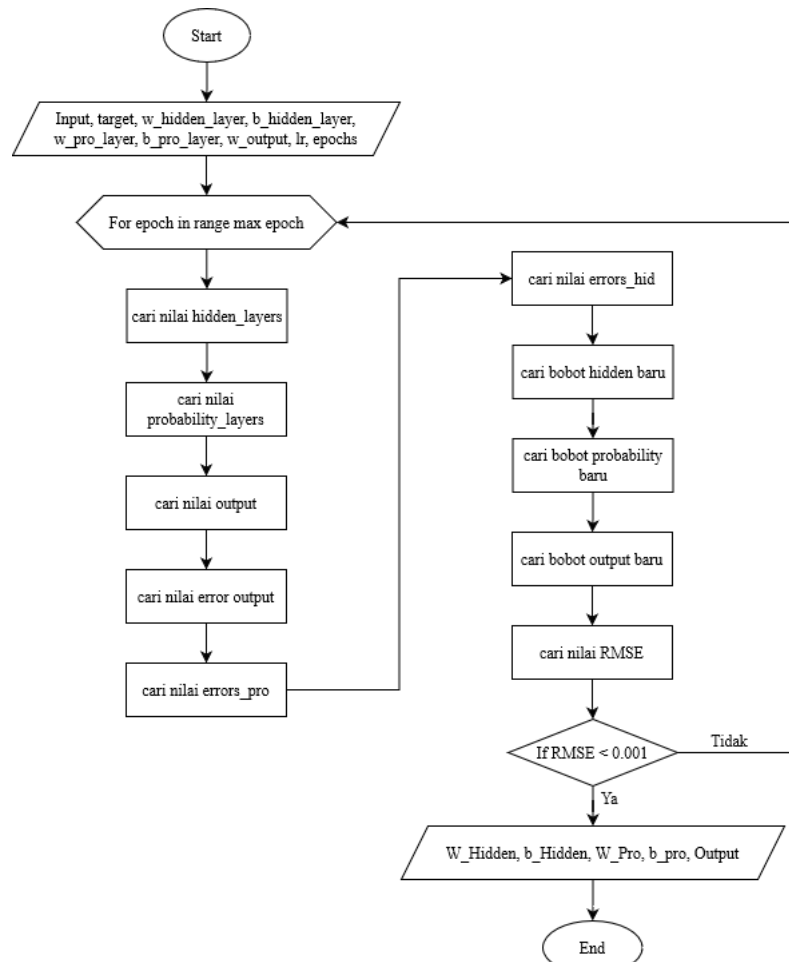
$$g_k = \text{sigmoid}\left(\sum_{j=1}^n h_j \cdot v_{jk} + b_j\right) \quad (3.4)$$

$$\text{output} = \text{sigmoid}\left(\sum_{k=1}^n g_k \cdot z_{k1}\right) \quad (3.5)$$

Keterangan,

- x_i : node indeks ke-i, berisi frekuensi kemunculan kata
- w_i : nilai antara *layer input* dengan *layer hidden*
- b_i : bias yang ada di *layer hidden*
- h_j : neuron yang ada di *layer hidden*

- v_j : nilai antara *layer hidden* dengan *layer probability output*
 b_j : bias yang ada di *layer probability output*
 g_k : *neuron* yang ada di *layer probability output*



Gambar 3.10 Flowchart Proses Training ANN

Pada Gambar 3.10 merupakan alur dari proses *training* ANN yang akan diimplementasikan ke dalam sebuah kode program, x_i diimplementasikan sebagai input, w_i diimplementasikan sebagai *w_hidden_layer*, b_i diimplementasikan sebagai *b_hidden_layer*, v_j diimplementasikan sebagai *w_pro_layer*, b_j diimplementasikan sebagai *b_pro_layer*, z_k diimplementasikan sebagai *w_output*. Gambar 3.11 hingga Gambar 3.13 merupakan implementasi proses *Artificial*

Neural Network pada sebuah *source code*. Gambar 3.14 merupakan *flowchart* proses *training*. Pada Gambar 3.15 merupakan *source code* dari fungsi prediksi.

```

1  import numpy as np
2
3  def sigmoid(x):
4      return 1/(1+np.exp(-x))
5
6  def training(input, target, w_hidden_layer, b_hidden_layer, w_pro_layer, b_pro_layer, w_output, lr,
7      epochs):
8
9      # inisialisasi jumlah
10     jml_input = len(input[0])
11     jml_hidden_layer = len(w_hidden_layer)
12     jml_propability_layer = len(w_pro_layer)
13     jml_target = len(target)
14
15     MSE = []
16     ACC = []
17     hasil_training = []
18
19     for epoch in range(epochs):
20         mse = 0
21         acc = 0
22         for idx in range(len(input)):
23             # feed forward
24             hidden_layers = []
25             propability_layers = []
26             #hidden layer
27             for h in range(jml_hidden_layer):
28                 hidden_layer = 0
29                 for n in range(jml_input):
30                     hidden_layer = hidden_layer+(input[idx][n]*w_hidden_layer[h][n])
31
32                 hidden_layer = hidden_layer+b_hidden_layer[h]
33                 hidden_layer = sigmoid(hidden_layer)
34                 hidden_layers.append(hidden_layer)
35             #probability layer
36             for p in range(jml_propability_layer):
37                 propability_layer = 0
38                 for hl, hidden_ly in enumerate(hidden_layers):
39                     propability_layer = propability_layer+(hidden_ly*w_pro_layer[p][hl])
40
41                 propability_layer = propability_layer+b_pro_layer[p]
42                 propability_layer = sigmoid(propability_layer)
43                 propability_layers.append(propability_layer)
44
45             output = 0
46             for pro, pro_ly in enumerate(propability_layers):
47                 output = output+(pro_ly*w_output[pro])
48             output = sigmoid(output)
49
50             # error
51             error = target[idx]-output
52
53             errors_pro = []
54             for w_o in w_output:
55                 error_pro = error*w_o
56                 errors_pro.append(error_pro)
57
58             errors_hid = []
59             for ids, err_pro in enumerate(errors_pro):
60                 error_hid = 0
61                 for p in range(jml_propability_layer):
62                     error_hid = error_hid+(w_pro_layer[ids][p]*err_pro)
63                 errors_hid.append(error_hid)
64

```

Gambar 3.11 *Source Code*-Proses ANN

```

67     # backpropagation
68     # bobot hidden baru
69     for h in range(jml_hidden_layer):
70         for i in range(jml_input):
71             w_hidden_layer[h][i] = w_hidden_layer[h][i]+lr*(errors_hid[h]*hidden_layers[h]*
72                 (1-hidden_layers[h])*input[idx][i])
73             b_hidden_layer[h]=b_hidden_layer[h]+lr*(errors_hid[h]*hidden_layers[h]*(1-hidden_layers
74                 [h]))
75     # bobot probability baru
76     for p in range(jml_propability_layer):
77         for h in range(jml_hidden_layer):
78             w_pro_layer[p][h] = w_pro_layer[p][h]+lr*(errors_pro[p]*propability_layers[p]*
79                 (1-propability_layers[p]))*hidden_layers[h]
80             b_pro_layer[p] = b_pro_layer[p]+lr*(errors_pro[p]*propability_layers[p]*
81                 (1-propability_layers[p]))
82     # bobot output baru
83     for p in range(jml_propability_layer):
84         w_output[p] = w_output[p]+lr*(error*output*(1-output))*propability_layers [p]
85     # mencari MSE dan ACC
86     mse = mse + error*error
87     acc = acc + np.absolute(target[idx]-output)
88
89     mse = mse/len(input)
90     acc = 1 - acc/len(input)
91
92     print("Epoch ",epoch)
93     print("MSE : ",mse)
94     print("ACC : ",acc)
95     print("=====")
96
97     MSE.append(mse)
98     ACC.append(acc)
99
100    if(mse < 0.001):
101        hasil_training.append(w_hidden_layer) #index 0 bobot hidden layer
102        hasil_training.append(b_hidden_layer) #index 1 bias hidden layer
103        hasil_training.append(w_pro_layer) #index 2 bobot probability layer
104        hasil_training.append(b_pro_layer) #index 3 bias probability layer
105        hasil_training.append(w_output) #index 4 bobot output layer
106        hasil_training.append(MSE) #index 5 MSE
107        hasil_training.append(ACC) #index 6 ACC
108
109    return hasil_training
110
111    hasil_training.append(w_hidden_layer) #index 0 bobot hidden layer
112    hasil_training.append(b_hidden_layer) #index 1 bias hidden layer
113    hasil_training.append(w_pro_layer) #index 2 bobot probability layer
114    hasil_training.append(b_pro_layer) #index 3 bias probability layer
115    hasil_training.append(w_output) #index 4 bobot output layer
116    hasil_training.append(w_output) #index 5 MSE
117    hasil_training.append(w_output) #index 6 ACC
118
119    return hasil_training

```

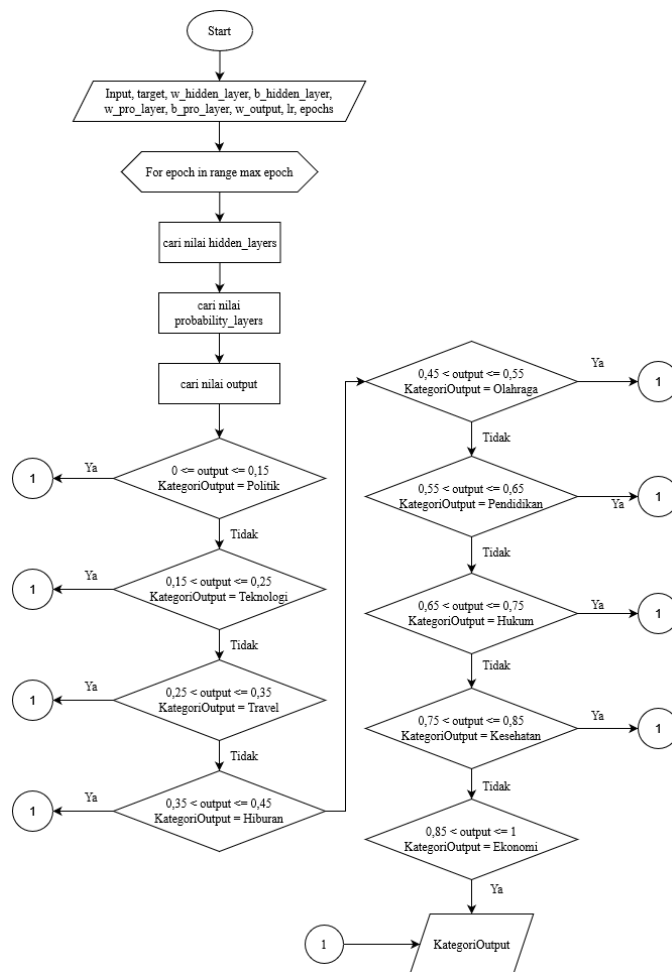
Gambar 3.12 Source Code- Proses ANN

```

1  import numpy as np
2
3  from data_train_90 import dt_train_90
4  from TF_Buffers import Get_TF_Buffers
5  from training import training
6  from Bobot import Get_Bobot
7
8  dt_train, data_hasils = dt_train_90()
9  TF_Buffers, target = Get_TF_Buffers(dt_train, data_hasils)
10
11 epochs = 2000
12 lr = 0.001
13 # W_Hidden, b_Hidden, W_Pro, b_Pro, W_Output = Get_Bobot()
14 jml_node_input = 500
15 jml_hidden_layer = 9
16 jml_pro_layer = 9
17 jml_out_layer = 9
18 W_Hidden = np.random.rand(jml_hidden_layer, jml_node_input) - 3
19 b_Hidden = np.random.rand(jml_hidden_layer) - 3
20 W_Pro = np.random.rand(jml_pro_layer, jml_hidden_layer) - 3
21 b_Pro = np.random.rand(jml_pro_layer) - 3
22 W_Output = np.random.rand(jml_out_layer) - 3
23 hasil_training = training(TF_Buffers, target, W_Hidden, b_Hidden, W_Pro, b_Pro, W_Output,
lr, epochs)

```

Gambar 3.13 Source Code- Proses ANN



Gambar 3.14 Flowchart Proses Testing ANN

```

39     if(dt_train[i]['Kategori'] == "Politik"):
40         target.append(0.1)
41     elif(dt_train[i]['Kategori'] == "Teknologi"):
42         target.append(0.2)
43     elif(dt_train[i]['Kategori'] == "Travel"):
44         target.append(0.3)
45     elif(dt_train[i]['Kategori'] == "Hiburan"):
46         target.append(0.4)
47     elif(dt_train[i]['Kategori'] == "Olahraga"):
48         target.append(0.5)
49     elif(dt_train[i]['Kategori'] == "Pendidikan"):
50         target.append(0.6)
51     elif(dt_train[i]['Kategori'] == "Hukum"):
52         target.append(0.7)
53     elif(dt_train[i]['Kategori'] == "Kesehatan"):
54         target.append(0.8)
55     elif(dt_train[i]['Kategori'] == "Ekonomi"):
56         target.append(0.9)
57
58     return TF_Buffers, target

```

Gambar 3.15 Source Code-Fungsi Prediksi

3.2.4.1 Training Artificial Neural Network

Pada tahap *training* akan dilakukannya penerapan pada metode *Neural Network*. Dengan data yang telah melalui proses tahap *Pre processing*, *Term frequency*, dan *Buffer Term*. Pada Tabel 3.5 merupakan beberapa data yang digunakan untuk proses *training*, daftar lengkap dari *training* terdapat pada Lampiran II.

Tabel 3.5 Data *Training*

No.	Judul Berita	Kategori
1.	Besok, Webinar Partai Perindo Bongkar Habis soal Moral Hazard Politik Uang di Level Voters	Politik
2.	Berstatus Tahanan Kota, Habib Rizieq: Tak Kurangi Semangat Selamatkan Indonesia	Politik
3.	Gunakan Sistem Baru, LYTO Luncukan Game Dekaron MMORPG	Teknologi
4.	Cara Mencari Masjid Terdekat Menggunakan Google Maps	Teknologi
5.	UGM dan Politeknik Pertanian Siap Lahirkan Petani Modern dengan Inovasi Terbaru	Pendidikan
6.	Galang Donasi Beasiswa, Unika Atma Jaya Gelar Lomba Lari	Pendidikan
7.	KPK Tetapkan Mantan Kepala BPKAD Jatim Tersangka Korupsi Bantuan Keuangan	Hukum
8.	Komisi III DPR Minta Satgassus Polri Diaudit	Hukum
9.	Tiket Pesawat Masih Mahal, Sandiaga Prediksi Tarif Turun Akhir Tahun	Travel
10.	Promo Tiket Garuda Indonesia, Jakarta-Labuan Bajo PP Rp 1,7 Jutaan	Travel
11.	Kenali Apa itu Niacinamide, Kandungan, Fungsi, dan Efek Sampinya	Kesehatan
12.	Apa Itu Batu Empedu?	Kesehatan
13.	Top Gun Maverick Geser Capaian Box Office Titanic di Amerika	Hiburan
14.	Review Manga: One Piece 1.056	Hiburan
15.	Kalteng Siap Gelar Piala Dunia MTB 2022	Olahraga
16.	Anthony Ginting di Kejuaraan Dunia 2022: Semoga Ketemu Axelsen	Olahraga

Pada tahap latih, *error* pada *epoch* yang didapatkan akan besar. Untuk mengecilkan *error* pada tiap *epoch* tersebut diperlukan pengoptimalan nilai bobot w . Output atau hasil prediksi yang didapatkan akan dihitung nilai *error*. Algoritma yang digunakan dalam pengoptimalan nilai bobot merupakan algoritma *backpropagation*. Nilai *error* akan dihitung dengan rumus RMSE (*Root Mean Square Error*) yang terdapat pada Persamaan 3.7:

$$error = \sqrt{(prediksi - target)^2} \quad (3.7)$$

Prediksi ialah *output* yang telah diprediksi oleh sebuah sistem, sedang target ialah hasil tujuan yang ingin diraih. *Error* didapatkan dengan cara melakukan perhitungan dengan rumus yang terdapat pada Persamaan 3.7. Selanjutnya untuk memperbarui bobot akan digunakan *Gradient Descent* yaitu menghitung selisih

antara nilai prediksi dari system dan nilai tagret sehingga nilai dari *gradient* akan didapatkan. Rumus *gradient descent* terdapat pada Persamaan 3.8:

$$gradient = (prediksi - target) . x_i \quad (3.9)$$

Iterasi yang digunakan agar mendapat nilai optimal dari parameter digambarkan sebagai *Gradient Descent*. Pada *Neural Network* khususnya pada bagian *backpropagation*, agar dapat memperbarui nilai bobot yaitu dengan meminimalkan *loss function* maka digunakanlah *Gradient Descent*. Rumus yang digunakan untuk memperbarui nilai bobot terdapat pada Persamaan 3.10:

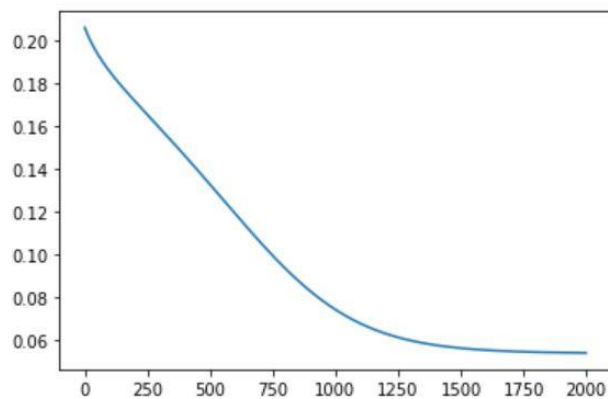
$$\theta_b = \theta_l - a.gradient \quad (3.10)$$

Dimana,

θ_b : nilai bobot yang baru

θ_l : nilai bobot yang lama

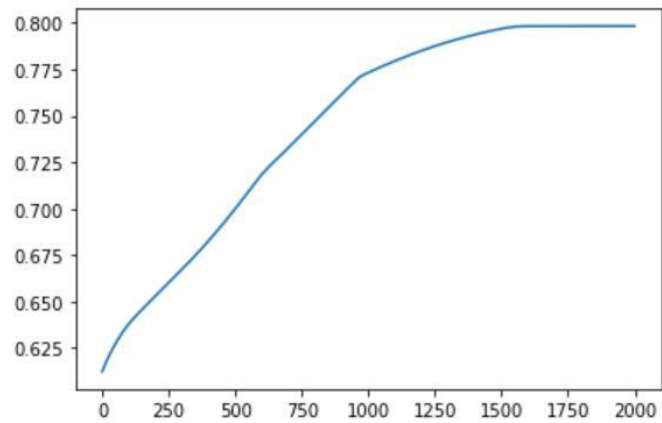
a : *learning rate*



Gambar 3.16 Grafik RMSE pada tahap *training*

Gambar 3.17 merupakan hasil grafik RMSE pada proses *training* pertama menggunakan rumus Persamaan 3.5, pada proses *training* di inialisasikan *learning rate* sebesar 0,001 dengan jumlah *neuron layer hidden* 9, jumlah

probability output 9, serta dengan nilai maksimal dari *epoch* ialah 2000. Dengan nilai *error* akhir sebesar 0.0539151331654491. Hasil *accuracy* pada proses *training* pertama dapat dilihat pada Gambar 3.18.



Gambar 3.17 Grafik *Accuracy* pada tahap *training*

BAB IV

UJI COBA DAN PEMBAHASAN

4.1 Skenario Pengujian

Skenario pengujian ialah skenario yang dilakukan untuk menguji sistem yang telah dibangun sebelumnya. Pada penelitian ini pengujian yang dilakukan ialah menghitung nilai *accuracy*, *precision*, *recall*, serta *f-measure*. Untuk pembagian dataset akan dilakukan dengan menggunakan tiga skenario pembagian antara data *training* dan data *testing*. Pembagian dataset ini dilakukan agar dapat menghasilkan perbandingan dari ketiga hasil klasifikasi yang akan dilakukan. Dari tiga skenario pembagian data tersebut dapat dilihat pada Tabel 4.1.

Tabel 4.1 Pembagian Dataset

Data Training	Data Testing
90%	10%
80%	20%
70%	30%

Hal pertama yang dilakukan ialah menemukan *ground truth*. *Ground truth* digunakan sebagai perbandingan hasil prediksi yang telah didapatkan sistem. Data yang digunakan pada penelitian ini merupakan data artikel berita berbahasa Indonesia yang berjumlah 360 data dengan sembilan kategori berita, dimana masing-masing kategori terdapat 40 artikel berita. Data berita dapat dilihat pada Lampiran III. *Ground truth* yang digunakan diambil dari sembilan kategori berita, jadi dapat dikatakan bahwa artikel yang sudah dikumpulkan bisa diklasifikasikan berdasarkan sembilan kategori berita, diantaranya politik, teknologi, travel, hiburan, olahraga, pendidikan, ekonomi, hukum, dan kesehatan. Nama dan

kategori yang dijadikan *ground truth* didapatkan berdasarkan kategori-kategori yang ada pada situs web penyedia berita online, yaitu detik.com, kompas.com, cnnindonesia.com, sindonews.com, dan juga republika.co.id dimana semua data tersebut dikumpulkan.

Pada pengujian pertama disiapkan data sebanyak 360 artikel berita, kemudian data dibagi antara data *training* dan data *testing* menjadi 90:10 didapatkan 324 artikel berita untuk data *training* dan 36 artikel berita untuk data *testing*. Apabila hasil telah didapatkan selanjutnya dihitung nilai *accuracy*, *precision*, *recall*, serta *f-measure*.

Selanjutnya pengujian kedua disiapkan data sebanyak 360 artikel berita, kemudian data dibagi antara data *training* dan data *testing* menjadi 80:20 didapatkan 288 artikel berita untuk data *training* dan 72 artikel berita untuk data *testing*. Apabila hasil telah didapatkan selanjutnya dihitung nilai *accuracy*, *precision*, *recall*, serta *f-measure*.

Kemudian pengujian ketiga disiapkan data sebanyak 360 artikel berita, kemudian data dibagi antara data *training* dan data *testing* menjadi 70:30 didapatkan 252 artikel berita untuk data *training* dan 108 artikel berita untuk data *testing*. Apabila hasil telah didapatkan selanjutnya dihitung nilai *accuracy*, *precision*, *recall*, serta *f-measure*.

Rumus yang digunakan untuk menghitung nilai *accuracy*, *precision*, *recall*, serta *f-measure* terdapat pada Persamaan 4.1 sampai Persamaan 4.4.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (4.1)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (4.2)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (4.3)$$

$$f - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4.4)$$

Keterangan,

1. TP (*True Positive*) : artikel dengan hasil pengklasifikasian kategori *i*, diprediksi benar oleh sistem bahwa kategori *i* masuk ke kategori *i*.
2. TN (*True Negative*) : artikel dengan hasil pengklasifikasian bukan kategori *i*, diprediksi benar oleh sistem bahwa kategori *i* tidak masuk ke kategori *i*.
3. FP (*False Positive*) : artikel dengan hasil pengklasifikasian bukan kategori *i*, namun diprediksi benar oleh sistem bahwa kategori *i* masuk ke kategori *i*.
4. FN (*False Negative*) : artikel dengan hasil pengklasifikasian kategori *i*, namun diprediksi salah oleh sistem bahwa kategori *i* tidak masuk ke kategori *i*.

4.2 Uji Coba

Berdasarkan skenario pengujian yang sudah dijabarkan pada sub-bab sebelumnya sistem akan melakukan prediksi terhadap berita yang sudah di *input* kan. Hasil prediksi dari sistem nantinya akan dibandingkan dengan *ground truth*. Pada pengujian pertama data yang digunakan sebanyak 324 artikel untuk dijadikan data *training* dan 36 artikel untuk dijadikan data *testing*. Hasil klasifikasi pengujian pertama dapat dilihat pada Tabel 4.2, dan dapat dilihat keseluruhan datanya di Lampiran IV. Tabel *confusion matrix* untuk pengujian 90:10 dapat dilihat pada Tabel 4.3.

Tabel 4.2 Hasil Klasifikasi Pengujian 90:10

Kategori/Prediksi	1	2	3	4	5	6	7	8	9
Politik	4	0	0	0	0	0	0	0	0
Teknologi	0	2	0	0	0	0	0	2	0
Travel	0	0	4	0	0	0	0	0	0
Hiburan	0	0	0	3	1	0	0	0	0
Olahraga	0	1	0	0	3	0	0	0	0
Pendidikan	0	0	0	0	1	3	0	0	0
Ekonomi	0	0	0	0	0	0	3	0	1
Hukum	0	0	0	0	0	0	0	4	0
Kesehatan	0	0	0	0	0	0	0	0	4

Tabel 4.3 *Confusion Matrix* untuk Pengujian 90:10

Kategori	<i>Confusion Matrix</i>			
	TP	TN	FP	FN
Politik	4	32	0	0
Teknologi	2	30	2	2
Travel	4	32	0	0
Hiburan	3	31	1	1
Olahraga	3	31	1	1
Pendidikan	3	31	1	1
Ekonomi	3	31	1	1
Hukum	4	32	0	0
Kesehatan	4	32	0	0
Total	30	282	6	6

Perhitungan *accuracy*, *precision*, *recall*, serta *f-measure* dari pengujian 90:10 adalah sebagai berikut:

$$Accuracy = \frac{312}{324} \times 100\% = 96,29\%$$

$$Precision = \frac{30}{36} \times 100\% = 83,33\%$$

$$Recall = \frac{30}{36} \times 100\% = 83,33\%$$

$$f - measure = 2 \times \frac{83,33 \times 83,33}{83,33 + 83,33} = 83,33\%$$

Hasil akhir dari pengujian 90:10 yaitu *accuracy* 96,29%, *precision* 83,33%, *recall* 83,33%, serta *f-measure* 83,33%.

Selanjutnya pada pengujian kedua data yang digunakan sebanyak 288 artikel berita untuk data *training* dan 72 artikel untuk dijadikan data *testing*. Hasil klasifikasi pengujian ketiga dapat dilihat pada Tabel 4.4, dan dapat dilihat keseluruhan datanya di Lampiran V. Tabel *confusion matrix* untuk pengujian 80:20 dapat dilihat pada Tabel 4.5.

Tabel 4.4 Hasil Klasifikasi Pengujian 80:20

Kategori/Prediksi	1	2	3	4	5	6	7	8	9
Politik	6	0	0	0	0	2	0	0	0
Teknologi	0	5	0	0	2	0	0	1	0
Travel	0	0	8	0	0	0	0	0	0
Hiburan	0	0	0	4	3	0	0	1	0
Olahraga	0	0	0	0	5	3	0	0	0
Pendidikan	0	0	0	0	0	6	0	2	0
Ekonomi	0	0	0	0	0	0	5	1	2
Hukum	0	0	0	0	0	4	0	4	0
Kesehatan	0	0	0	0	0	0	0	3	5

Tabel 4.5 *Confusion Matrix* untuk Pengujian 80:20

Kategori	<i>Confusion Matrix</i>			
	TP	TN	FP	FN
Politik	6	62	2	2
Teknologi	5	61	3	3
Travel	8	64	0	0
Hiburan	4	60	4	4
Olahraga	5	61	3	3
Pendidikan	6	62	2	2
Ekonomi	5	61	3	3
Hukum	4	60	4	4
Kesehatan	5	61	3	3
Total	48	552	24	24

Perhitungan *accuracy*, *precision*, *recall*, serta *f-measure* dari pengujian 80:20 adalah sebagai berikut:

$$Accuracy = \frac{600}{648} \times 100\% = 92,59\%$$

$$Precision = \frac{48}{72} \times 100\% = 66,66\%$$

$$Recall = \frac{48}{72} \times 100\% = 66,66\%$$

$$f - measure = 2 \times \frac{66,66 \times 66,66}{66,66 + 66,66} = 66,66\%$$

Hasil akhir dari pengujian 80:20 yaitu *accuracy* 92,59%, *precision* 66,66%, *recall* 66,66%, serta *f-measure* 66,66%.

Selanjutnya pada pengujian ketiga data yang digunakan sebanyak 252 artikel berita untuk data *training* dan 108 artikel untuk dijadikan data *testing*. Hasil klasifikasi pengujian kelima dapat dilihat pada Tabel 4.6, dan dapat dilihat keseluruhan datanya di Lampiran VI. Tabel *confusion matrix* untuk pengujian 70:30 dapat dilihat pada Tabel 4.7.

Tabel 4.6 Hasil Klasifikasi Pengujian 70:30

Kategori/Prediksi	1	2	3	4	5	6	7	8	9
Politik	6	0	0	0	0	6	0	0	0
Teknologi	0	5	0	3	0	0	0	4	0
Travel	0	0	6	0	0	0	6	0	0
Hiburan	0	0	0	7	3	0	0	2	0
Olahraga	0	7	0	0	5	0	0	0	0
Pendidikan	0	0	0	4	0	6	0	2	0
Ekonomi	0	4	0	0	0	0	5	0	3
Hukum	3	0	0	0	0	4	0	5	0
Kesehatan	0	0	0	0	5	0	1	0	6

Tabel 4.7 *Confusion Matrix* untuk Pengujian 70:30

Kategori	<i>Confusion Matrix</i>			
	TP	TN	FP	FN
Politik	6	90	6	6
Teknologi	5	89	7	7
Travel	6	90	6	6
Hiburan	7	91	5	5
Olahraga	5	89	7	7
Pendidikan	6	90	6	6
Ekonomi	5	89	7	7
Hukum	5	89	7	7
Kesehatan	6	90	6	6
Total	51	807	57	57

Perhitungan *accuracy*, *precision*, *recall*, serta *f-measure* dari pengujian 70:30 adalah sebagai berikut:

$$Accuracy = \frac{858}{972} \times 100\% = 88,27\%$$

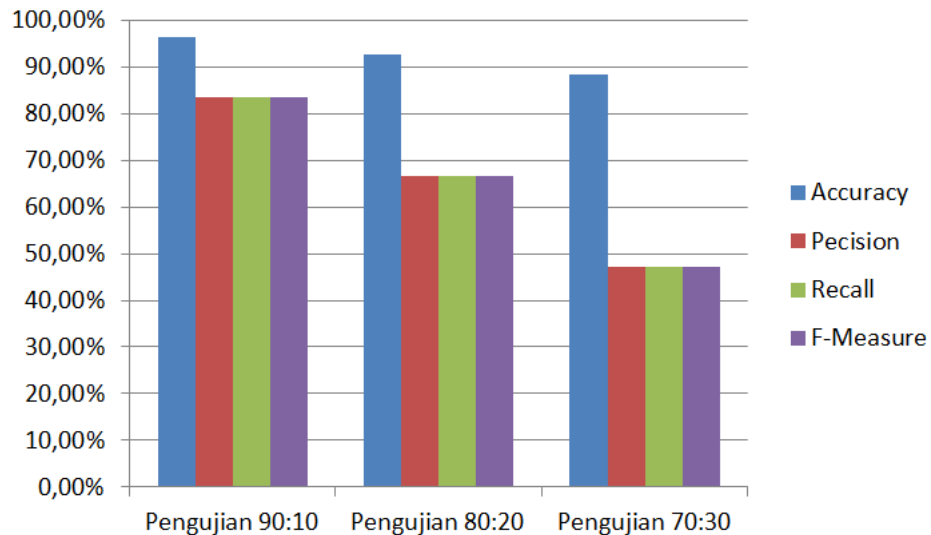
$$Precision = \frac{51}{108} \times 100\% = 47,22\%$$

$$Recall = \frac{51}{108} \times 100\% = 47,22\%$$

$$f - measure = 2 \times \frac{47,22 \times 47,22}{47,22 + 47,22} = 47,22\%$$

Hasil akhir dari pengujian 70:30 yaitu *accuracy* 88,27%, *precision* 47,22%, *recall* 47,22%, serta *f-measure* 47,22%.

4.3 Pembahasan



Gambar 4.1 Gambar Perbandingan

Hasil perbandingan dari ketiga pengujian dapat dilihat pada Gambar 4.1. Hasil akhir dari pengujian 90:10 menggunakan rumus Persamaan 3.5 yaitu *accuracy* 96,29%, *precision* 83,33%, *recall* 83,33%, serta *f-measure* 83,33%. Hasil dari pengujian 80:20 yaitu *accuracy* 92,59%, *precision* 66,66%, *recall* 66,66%, serta *f-measure* 66,66%. Hasil dari pengujian 70:30 yaitu *accuracy* 88,27%, *precision* 47,22%, *recall* 47,22%, serta *f-measure* 47,22%.

Berdasarkan hasil uji coba jumlah data yang digunakan pada saat proses *training* sangat berpengaruh terhadap proses pengujian, dimana semakin besar data yang digunakan untuk *training* maka semakin baik sistem dalam melakukan proses pengujian, sebaliknya semakin sedikit data yang digunakan pada proses *training* maka semakin rendah hasil *accuracy* yang didapatkan oleh sistem.

Saat proses prediksi, bobot sangat berpengaruh terhadap *Neural Network*, maka dari itu pada saat dilakukannya proses pembelajaran nilai bobot akan diperbarui menggunakan selisih dari hasil prediksi yang didapatkan oleh sistem dan *ground truth* sehingga mendapatkan bobot yang optimal. Dalam proses pembelajaran *learning rate* serta *epoch* sangat berpengaruh terhadap hasil yang didapat, apabila *learning rate* tinggi maka dapat menyebabkan *overfitting model*. Selain *learning rate* dan bobot, data yang di *input* kan juga berpengaruh terhadap *Neural Network*, karena hasil prediksi dari *Neural Network* didapatkan berdasarkan pola dari *input* an, apabila data yang di *input* kan tidak jelas, maka dapat menyebabkan kesalahan pada proses prediksi.

Penerapan *mu'amalah ma'a Allah* pada penelitian ini yaitu, dengan adanya klasifikasi berita online manusia diberi kemampuan untuk mengembangkan teknologi sehingga memudahkan manusia dalam melakukan pekerjaan. Diharapkan penelitian ini dapat memudahkan masyarakat dalam mencari tahu keadaan yang terjadi saat ini. Dimana masyarakat hanya perlu mencari berita sesuai dengan kategori yang diinginkan agar dapat mengetahui apa saja yang terjadi disekitarnya, misalnya berita tentang fenomena alam, perkembangan teknologi, serta berita-berita lainnya. Melalui penelitian ini sekiranya dapat membantu untuk memudahkan para editor dalam bekerja serta memudahkan masyarakat dalam mengakses berita, sehingga masyarakat dapat selalu mengingat Allah *Subhanahu wa ta'ala* serta bersyukur dengan apa yang sudah diberikan Allah *Subhanahu wa ta'ala*. Pada Qur'an Surah Al-A'raf ayat 56 Allah *Subhanahu wa ta'ala* berfirman:

وَلَا تُفْسِدُوا فِي الْأَرْضِ بَعْدَ إِصْلَاحِهَا وَادْعُوهُ خَوْفًا وَطَمَعًا ۚ إِنَّ رَحْمَتَ اللَّهِ قَرِيبٌ مِّنَ الْمُحْسِنِينَ

“Dan janganlah kamu berbuat kerusakan di bumi setelah (diciptakan) dengan baik. Berdoalah kepada-Nya dengan rasa takut dan penuh harap. Sesungguhnya rahmat Allah sangat dekat kepada orang yang berbuat kebaikan.” (QS. Al-A’raf 7:56)

Pada *Zubdatut Tafsir Min Fathil Qadir* (TafsirWeb, n.d.) menjelaskan “Dalam ayat إِنَّ رَحْمَتَ اللَّهِ قَرِيبٌ مِّنَ الْمُحْسِنِينَ terdapat dorongan dan pemberian semangat bagi hamba-hamba-Nya untuk berbuat kebaikan. Dan (المحسنون) yaitu orang-orang yang mengumpulkan dalam dirinya keimanan kepada Allah dan kepada hal yang ghaib, dan menjalankan kewajiban-kewajiban kepada Allah, dan menjauhi larangan-larangan-Nya, serta merasa diawasi oleh Allah sehingga mereka memperbaiki amalan mereka”. Pada ayat tersebut Allah *Subhanahu wa ta’ala* menjelaskan bahwa rahmat Allah sangat dekat kepada orang yang berbuat kebaikan, diharapkan penelitian ini dapat bermanfaat sebagaimana semestinya serta dapat memberi kebaikan bagi para penggunanya dalam mengelompokan berita-berita berdasarkan dengan kategori-kategori yang benar.

Selanjutnya yaitu penerapan *mu’amalah ma’a An-nas* pada penelitian ini yaitu, memudahkan para editor dalam mengklasifikasikan berita sehingga para editor tidak perlu mengklasifikasikan berita secara manual serta memudahkan para pembaca dalam mengakses berita yang ingin dibaca sehingga para pembaca dapat dengan mudah membaca berita sesuai dengan keinginannya. Pembaca hanya perlu mencari kategori berita yang ingin dibaca, sehingga tidak perlu mengkhawatirkan apakah berita itu sudah sesuai dengan kategori atau tidak. Oleh

karena itu diharapkan penelitian ini dapat memberikan kemudahan bagi para editor dalam mengklasifikasikan berita serta memberikan kemudahan bagi masyarakat dalam mencari berita sesuai dengan kategori berita. Dalam Qur'an surah Al-Hujurat Allah *Subhanahu wa ta'ala* berfirman:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِن جَاءَكُمْ فَاسِقٌ بِنَبَأٍ فَتَبَيَّنُوا أَن تُصِيبُوا قَوْمًا بِجَهَالَةٍ فَتُصْحَبُوا عَلَىٰ مَا فَعَلْتُمْ نَادِمِينَ

“Hai orang-orang yang beriman, jika datang kepadamu orang fasik membawa suatu berita, maka periksalah dengan teliti, agar kamu tidak menimpakan suatu musibah kepada suatu kaum tanpa mengetahui keadaannya yang menyebabkan kamu menyesal atas perbuatanmu itu“. (QS. Al-Hujurat 49:6)

Pada tafsir *Al-Muyassar* (TafsirWeb, n.d.) dijelaskan “Wahai orang-orang yang membenarkan Allah dan RasulNya serta melaksanakan syariatNya, bila orang fasik datang kepada kalian dengan membawa sebuah berita, maka periksalah beritanya sebelum membenarkan dan menukilnya agar kalian mengetahui kebenarannya, di khawatirkan kalian bisa melakukan tindakan zhalim terhadap suatu kaum yang tidak bersalah, akibatnya kalian akan menyesalinya.”

Pada ayat *يَا أَيُّهَا الَّذِينَ آمَنُوا إِن جَاءَكُمْ فَاسِقٌ بِنَبَأٍ فَتَبَيَّنُوا* Allah telah menjelaskan “Hai orang-orang yang beriman, jika datang kepadamu orang fasik membawa suatu berita, maka periksalah dengan teliti”, maksud dari ayat tersebut ialah apabila kita menerima suatu berita kita harus mencari tahu lebih dalam kebenaran berita tersebut. Diharapkannya sistem klasifikasi berita online ini bermanfaat bagi para editor dalam melakukan tugasnya, yakni mengelompokan berita-berita berdasarkan dengan kategori-kategori yang benar sehingga masyarakat dapat dengan mudah mengakses berita berdasarkan kategori yang ingin dibaca.

BAB V

PENUTUP

5.1 Kesimpulan

Setelah dilakukannya pengujian terhadap pengklasifikasian artikel berita online menggunakan metode *Artificial Neural Network*, dengan menggunakan 3 pengujian data dapat disimpulkan bahwa.

1. Hasil dari pengujian 90:10 yaitu *accuracy* 96,29%, *precision* 83,33%, *recall* 83,33%, serta *f-measure* 83,33%. Hasil dari pengujian 80:20 yaitu *accuracy* 92,59%, *precision* 66,66%, *recall* 66,66%, serta *f-measure* 66,66%. Hasil dari pengujian 70:30 yaitu *accuracy* 88,27%, *precision* 47,22%, *recall* 47,22%, serta *f-measure* 47,22%.
2. Faktor-faktor yang mempengaruhi performa pada klasifikasi berita online menggunakan metode *Artificial Neural Network* ialah jumlah data yang digunakan pada saat proses *training* sangat berpengaruh terhadap hasil akhir, sehingga semakin banyak data yang digunakan pada proses *training* maka semakin baik pula hasil yang didapatkan. Hal ini disebabkan adanya pola pembelajaran pada sebuah sistem.

5.2 Saran

Setelah dilakukannya pengujian terhadap pengklasifikasian artikel berita online menggunakan metode *Artificial Neural Network*. Peneliti menyadari bahwa sistem yang dibangun belumlah sempurna, sehingga agar sistem dibangun jauh

lebih baik dibutuhkannya pengembangan terhadap sistem. Berikut merupakan saran kepada para peneliti selanjutnya:

1. Menggunakan metode *feature selection* yang lain seperti halnya *Information Gain*, *Mutual Information*, *Fast Correlation Based Filter*, dan *feature selection* lainnya.
2. Pada proses pengujian agar dapat menerapkan proses validasi model dengan menggunakan metode *K-Fold Cross Validation*, serta menggunakan *Early Stopping* pada proses *training model* agar dapat menghindari *overfitting model* saat melakukan proses *training model*.
3. Akan lebih baik apabila jumlah data yang digunakan lebih besar karena berpengaruh terhadap kinerja sistem, serta data yang digunakan tidak hanya berita berbahasa Indonesia.

DAFTAR PUSTAKA

- Andika, L. A., Pratiwi, H., & Handajani, S. S. (2019). Lingga Aji Andika 1 , Hasih Pratiwi 2 , and Sri Sulistijowati Handajani 3 1. *Indonesian Journal of Statistics and Its Applications*, 3(3), 331–340.
- Atiliani. (2013). *Pelatihan Jaringan Syaraf Tiruan Multi Layer Perceptron Menggunakan Genetic Algorithm Levenberg Marquardt*.
- Crowston, K., Allen, E. E., & Heckman, R. (2012). Using natural language processing technology for qualitative data analysis. *International Journal of Social Research Methodology*, 15(6), 523–543. <https://doi.org/10.1080/13645579.2011.625764>
- Dio Ariadi, & Kartika Fithriasari. (2015). Klasifikasi Berita Indonesia Menggunakan Metode Naive Bayesian Classification dan Support Vector Machine dengan Confix Stripping Stemmer. *Jurnal Sains Dan Seni ITS*, 4(2).
- Hanifa, A., Fauzan, S. A., Hikal, M., & ... (2021). Perbandingan Metode LSTM dan GRU (RNN) untuk Klasifikasi Berita Palsu Berbahasa Indonesia. *Dinamika Rekayasa*, 17(1), 33–39.
- Herwijayanti, B., Ratnawati, D. E., & Muflikhah, L. (2018). Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity. *Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(1), 306–312.
- Hidayat, E. Y., & Rizqi, M. A. (2020). Klasifikasi Dokumen Berita Menggunakan Algoritma Enhanced Confix Stripping Stemmer dan Naïve Bayes Classifier. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 6(2), 90–99. <https://doi.org/10.25077/teknosi.v6i2.2020.90-99>
- Indah, S., & Setiawan, A. (2011). Penerapan Jaringan Saraf Tiruan Metode Backpropagation Menggunakan VB 6. *Ultimatics*, III(2), 23.
- Muktafin, E. H., Kusriani, K., & Luthfi, E. T. (2020). Analisis Sentimen pada Ulasan Pembelian Produk di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing. *Jurnal Eksplora Informatika*, 10(1), 32–42. <https://doi.org/10.30864/eksplora.v10i1.390>
- Munasatya, N., & Novianto, S. (2020). Natural Language Processing untuk Sentimen Analisis Presiden Jokowi Menggunakan Multi Layer Perceptron. *Techno.Com*, 19(3), 237–244. <https://doi.org/10.33633/tc.v19i3.3630>
- Nawawi, H. M., Purnama, J. J., & Hikmah, A. B. (2019). Komparasi Algoritma Neural Network Dan Naïve Bayes Untuk Memprediksi Penyakit Jantung. *Jurnal Pilar Nusa Mandiri*, 15(2), 189–194. <https://doi.org/10.33480/pilar.v15i2.669>
- Prasetyo, V. R., Benarkah, N., & Chrisintha, V. J. (2021). Implementasi Natural Language Processing Dalam Pembuatan Chatbot Pada Program Information Technology Universitas Surabaya. *Teknika*, 10(2), 114–121. <https://doi.org/10.34148/teknika.v10i2.370>
- Rhomadhona, H., & Permadi, J. (2019). Klasifikasi Berita Kriminal

- Menggunakan Naïve Bayes Classifier (NBC) dengan Pengujian K-Fold Cross Validation. *Jurnal Sains Dan Informatika*, 5(2), 108–117. <https://doi.org/10.34128/jsi.v5i2.177>
- Rini, D. C., Farida, Y., & Puspitasari, D. (2016). Klasifikasi Menggunakan Metode Hybrid Bayesian-Neural Network (Studi Kasus: Identifikasi Virus Komputer). *Jurnal Matematika "MANTIK,"* 1(2), 38. <https://doi.org/10.15642/mantik.2016.1.2.38-43>
- Risdianto. (2021). *Klasifikasi Dokumen Perundang-Undangan Menggunakan Metode Jaringan Saraf Tiruan.*
- Rohman, A. N., Utami, E., & Raharjo, S. (2019). Deteksi Kondisi Emosi pada Media Sosial Menggunakan Pendekatan Leksikon dan Natural Language Processing. *Eksplora Informatika*, 9(1), 70–76. <https://doi.org/10.30864/eksplora.v9i1.277>
- Subroto, G. (2019). Klasifikasi bahan pustaka. *Pustakawan Perpustakaan UM, Ddc*, 1–13.
- Tala, F. Z. (2013). A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia. *M.Sc. Thesis, Appendix D, pp*, 39–46.
- Wangsanegara, Nisa Kurniasih, & Subaeki, B. (2015). Implementasi Natural Language Processing Dalam Pengukuran Ketepatan Ejaan Yang Disempurnakan (EYD) Pada Abstrak Skripsi Menggunakan Algoritma Fuzzy Logic. *Jurnal Teknik Informatika*, 8(2).
- Wisnu Widiatmoko. (2013). Analisis Kohesi Dan Koherensi Wacana Berita Rubrik Nasional Di Majalah Online Detik. *Jurnal Sastra Indonesia*, 2(1), 1–7.
- Wongso, R., Luwinda, F. A., Trisnajaya, B. C., Rusli, O., & Rudy. (2017). News Article Text Classification in Indonesian Language. *Procedia Computer Science*, 116, 137–143. <https://doi.org/10.1016/j.procs.2017.10.039>
- Zamani, A. M., & Amaliah, B. (2012). *Implementasi Algoritma Genetika pada Struktur Backpropagation Neural Network untuk Klasifikasi Kanker Payudara. 1.*

LAMPIRAN

Lampiran I

Daftar Stopword Removal

<i>Daftar Stopword</i>			
Ada	Dimulai	Melakukan	Sejak
Adalah	Dimulailah	Melalui	Sejauh
Adanya	Dimulainya	Melihat	Sejenak
Adapun	Dimungkinkan	Melihatnya	Sejumlah
Agak	Dini	Memang	Sekadar
Agaknya	Dipastikan	Memastikan	Sekadarnya
Agar	Diperbuat	Memberi	Sekali
Akan	Diperbuatnya	Memberikan	Sekali-kali
Akankah	Dipergunakan	Membuat	Sekalian
Akhir	Diperkirakan	Memerlukan	Sekaligus
Akhiri	Diperlihatkan	Memihak	Sekalipun
Akhirnya	Diperlukan	Meminta	Sekarang
Aku	Diperlukannya	Memintakan	Sekarang
Akulah	Dipersoalkan	Memisalkan	Sekecil
Amat	Dipertanyakan	Memperbuat	Seketika
Amatlah	Dipunyai	Mempergunakan	Sekiranya
Anda	Diri	Memperkirakan	Sekitar
Andalah	Dirinya	Memperlihatkan	Sekitarnya
Antar	Disampaikan	Mempersiapkan	Sekurang-kurangnya
Antara	Disebut	Mempersoalkan	Sekurangnya
Antaranya	Disebutkan	Mempertanyakan	Sela
Apa	Disebutkannya	Mempunyai	Selain
Apaan	Disini	Memulai	Selaku
Apabila	Disinilah	Memungkinkan	Selalu
Apakah	Ditambahkan	Menaiki	Selama
Apalagi	Ditandakan	Menambahkan	Selama-lamanya
Apatah	Ditanya	Menandakan	Selamanya
Artinya	Ditanyai	Menanti	Selanjutnya
Asal	Ditanyakan	Menanti-nanti	Seluruh
Asalkan	Ditegaskan	Menantikan	Seluruhnya
Atas	Ditujukan	Menanya	Semacam
Atau	Ditunjuk	Menanyai	Semakin
Ataukah	Ditunjuki	Menanyakan	Semampu
Ataupun	Ditunjukkan	Mendapat	Semampunya
Awal	Ditunjukkannya	Mendapatkan	Semasa
Awalnya	Ditunjuknya	Mendatang	Semasih
Bagai	Dituturkan	Mendatangi	Semata
Bagaikan	Dituturkannya	Mendatangkan	Semata-mata

Bagaimana	Diucapkan	Menegaskan	Semaunya
Bagaimanakah	Diucapkannya	Mengakhiri	Sementara
Bagaimanapun	Diungkapkan	Mengapa	Semisal
Bagi	Dong	Mengatakan	Semisalnya
Bagian	Dua	Mengatakannya	Sempat
Bahkan	Dulu	Mengenai	Semua
Bahwa	Empat	Mengerjakan	Semuanya
Bahwasanya	Enggak	Mengetahui	Semula
Baik	Enggaknya	Menggunakan	Sendiri
Bakal	Entah	Menghendaki	Sendirian
Bakalan	Entahlah	Mengibaratkan	Sendirinya
Balik	Guna	Mengibaratkannya	Seolah
Banyak	Gunakan	Mengingat	Seolah-olah
Bapak	Hal	Mengingatkan	Seorang
Baru	Hampir	Menginginkan	Sepanjang
Bawah	Hanya	Mengira	Sepantasnya
Beberapa	Hanyalah	Mengucapkan	Sepantasnyalah
Begini	Hari	Mengucapkannya	Seperlunya
Beginian	Harus	Mengungkapkan	Seperti
Beginikah	Haruslah	Menjadi	Sepertinya
Beginilah	Harusnya	Menjawab	Sepihak
Begitu	Hendak	Menjelaskan	Sering
Begitukah	Hendaklah	Menuju	Seringnya
Begitulah	Hendaknya	Menunjuk	Serta
Begitupun	Hingga	Menunjuki	Serupa
Bekerja	Ia	Menunjukkan	Sesaat
Belakang	Ialah	Menunjuknya	Sesama
Belakangan	Ibarat	Menurut	Sesampai
Belum	Ibaratkan	Menuturkan	Sesegera
Belumlah	Ibaratnya	Menyampaikan	Sesekali
Benar	Ibu	Menyangkut	Seseorang
Benarkah	Ikut	Menyatakan	Sesuatu
Benarlah	Ingat	Menyebutkan	Sesuatunya
Berada	Ingat-ingat	Menyeluruh	Sesudah
Berakhir	Ingin	Menyiapkan	Sesudahnya
Berakhirilah	Inginkah	Merasa	Setelah
Berakhirnya	Inginkan	Mereka	Setempat
Berapa	Ini	Merekalah	Setengah
Berapakah	Inikah	Merupakan	Seterusnya
Berapalah	Inilah	Meski	Setiap
Berapapun	Itu	Meskipun	Setiba
Berarti	Itukah	Meyakini	Setibanya

Berawal	Itulah	Meyakinkan	Setidak-tidaknya
Berbagai	Jadi	Minta	Setidaknya
Berdatangan	Jadilah	Mirip	Setinggi
Beri	Jadinya	Misal	Seusai
Berikan	Jangan	Misalkan	Sewaktu
Berikut	Jangankan	Misalnya	Siap
Berikutnya	Janganlah	Mula	Siapa
Berjumlah	Jauh	Mulai	Siapakah
Berkali-kali	Jawab	Mulailah	Siapapun
Berkata	Jawaban	Mulanya	Sini
Berkehendak	Jawabnya	Mungkin	Sinilah
Berkeinginan	Jelas	Mungkinkah	Soal
Berkenaan	Jelaskan	Nah	Soalnya
Berlainan	Jelaslah	Naik	Suatu
Berlalu	Jelasnya	Namun	Sudah
Berlangsung	Jika	Nanti	Sudahkah
Berlebihan	Jikalau	Nantinya	Sudahlah
Bermacam	Juga	Nyaris	Supaya
Bermacam-macam	Jumlah	Nyatanya	Tadi
Bermaksud	Jumlahnya	Oleh	Tadinya
Bermula	Justru	Olehnya	Tahu
Bersama	Kala	Pada	Tahun
Bersama-sama	Kalau	Padahal	Tak
Bersiap	Kalaulah	Padanya	Tambah
Bersiap-siap	Kalaupun	Pak	Tambahnya
Bertanya	Kalian	Paling	Tampak
Bertanya-tanya	Kami	Panjang	Tampaknya
Berturut	Kamilah	Pantas	Tandas
Berturut-turut	Kamu	Para	Tandasnya
Bertutur	Kamulah	Pasti	Tanpa
Berujar	Kan	Pastilah	Tanya
Berupa	Kapan	Penting	Tanyakan
Besar	Kapankah	Pentingnya	Tanyanya
Betul	Kapanpun	Per	Tapi
Betulkah	Karena	Percuma	Tegas
Biasa	Karenanya	Perlu	Tegasnya
Biasanya	Kasus	Perlukah	Telah
Bila	Kata	Perlunya	Tempat
Bilakah	Katakan	Pernah	Tengah
Bisa	Katakanlah	Persoalan	Tentang
Bisakah	Katanya	Pertama	Tentu
Boleh	Ke	Pertama-tama	Tentulah

Bolehkah	Keadaan	Pertanyaan	Tentunya
Bolehlah	Kebetulan	Pertanyakan	Tepat
Buat	Kecil	Pihak	Terakhir
Bukan	Kedua	Pihaknya	Terasa
Bukankah	Keduanya	Pukul	Terbanyak
Bukanlah	Keinginan	Pula	Terdahulu
Bukannya	Kelamaan	Pun	Terdapat
Bulan	Kelihatan	Punya	Terdiri
Bung	Kelihatannya	Rasa	Terhadap
Cara	Kelima	Rasanya	Terhadapnya
Caranya	Keluar	Rata	Teringat
Cukup	Kembali	Rupanya	Teringat-ingat
Cukupkah	Kemudian	Saat	Terjadi
Cukuplah	Kemungkinan	Saatnya	Terjadilah
Cuma	Kemungkinannya	Saja	Terjadinya
Dahulu	Kenapa	Sajalah	Terkira
Dalam	Kepada	Saling	Terlalu
Dan	Kepadanya	Sama	Terlebih
Dapat	Kesampaian	Sama-sama	Terlihat
Dari	Keseluruhan	Sambil	Termasuk
Daripada	Keseluruhannya	Sampai	Ternyata
Datang	Keterlalu	Sampai-sampai	Tersampaikan
Dekat	Ketika	Sampaikan	Tersebut
Demi	Khususnya	Sana	Tersebutlah
Demikian	Kini	Sangat	Tertentu
Demikianlah	Kinilah	Sangatlah	Tertuju
Dengan	Kira	Satu	Terus
Depan	Kira-kira	Saya	Terutama
Di	Kiranya	Sayalah	Tetap
Dia	Kita	Se	Tetapi
Diakhiri	Kitalah	Sebab	Tiap
Diakhirinya	Kok	Sebabnya	Tiba
Dialah	Kurang	Sebagai	Tiba-tiba
Diantara	Lagi	Sebagaimana	Tidak
Diantaranya	Lagian	Sebagainya	Tidakkah
Diberi	Lah	Sebagian	Tidaklah
Diberikan	Lain	Sebaik	Tiga
Diberikannya	Lainnya	Sebaik-baiknya	Tinggi
Dibuat	Lalu	Sebaiknya	Toh
Dibuatnya	Lama	Sebaliknya	Tunjuk
Didapat	Lamanya	Sebanyak	Turut
Didatangkan	Lanjut	Sebegini	Tutur

Digunakan	Lanjutnya	Sebegitu	Tuturnya
Diibaratkan	Lebih	Sebelum	Ucap
Diibaratkannya	Lewat	Sebelumnya	Ucapnya
Diingat	Lima	Sebenarnya	Ujar
Diingatkan	Luar	Seberapa	Ujarnya
Diinginkan	Macam	Sebesar	Umum
Dijawab	Maka	Sebetulnya	Umumnya
Dijelaskan	Makanya	Sebisanya	Ungkap
Dijelaskannya	Makin	Sebuah	Ungkapnya
Dikarenakan	Malah	Sebut	Untuk
Dikatakan	Malahan	Sebutlah	Usah
Dikatakannya	Mampu	Sebutnya	Usai
Dikerjakan	Mampukah	Secara	Waduh
Diketahui	Mana	Secukupnya	Wah
Diketuinya	Manakala	Sedang	Wahai
Dikira	Manalagi	Sedangkan	Waktu
Dilakukan	Masa	Sedemikian	Waktunya
Dilalui	Masalah	Sedikit	Walau
Dilihat	Masalahnya	Sedikitnya	Walaupun
Dimaksud	Masih	Seenaknya	Wong
Dimaksudkan	Masihkah	Segala	Yaitu
Dimaksudkannya	Masing	Segalanya	Yakin
Dimaksudnya	Masing-masing	Segeera	Yakni
Diminta	Mau	Seharusnya	Yang
Dimintai	Maupun	Sehingga	
Dimisalkan	Melainkan	Seingat	

Lampiran II

Daftar data *training* berita

No.	Judul Berita	Kategori
1	Besok, Webinar Partai Perindo Bongkar Habis soal Moral Hazard Politik Uang di Level Voters	Politik
2	Berstatus Tahanan Kota, Habib Rizieq: Tak Kurangi Semangat Selamatkan Indonesia	Politik
3	Perempuan PAN Didorong Persiapkan Diri Hadapi Pemilu 2024	Politik
4	Akan Ada Parpol Gabung KIB, Golkar: Tidak Mengubah dan Mengganggu Irama Koalisi	Politik
5	Ratusan Milenial di Depok Dukung Sandiaga Uno Maju di Pilpres 2024	Politik
6	Gunakan Sistem Baru, LYTO Luncurkan Game Dekaron MMORPG	Teknologi

7	Cara Mencari Masjid Terdekat Menggunakan Google Maps	Teknologi
8	Google Sempat Down Selama 1 Jam, Begini Dampaknya	Teknologi
9	5 Fakta Google Maps yang Belum Banyak Diketahui Orang	Teknologi
10	Samsung Akan Gunakan Teknologi UI Terbaru untuk Galaxy S22	Teknologi
11	UGM dan Politeknik Pertanian Siap Lahirkan Petani Modern dengan Inovasi Terbaru	Pendidikan
12	Galang Donasi Beasiswa, Unika Atma Jaya Gelar Lomba Lari	Pendidikan
13	Dorong Ekonomi Sirkular, Mahasiswa ITS Gagas Waralaba Pengolahan Kompos	Pendidikan
14	Cerita Inspiratif Salsabilla, Wisudawan Teman Tuli yang Sukses Lulus dari ITB	Pendidikan
15	Viral Curhatan Wanita yang Berjuang Dapatkan 4 Gelar Sekaligus dari UI, UGM, BINUS, dan USYD	Pendidikan
16	KPK Tetapkan Mantan Kepala BPKAD Jatim Tersangka Korupsi Bantuan Keuangan	Hukum
17	Komisi III DPR Minta Satgassus Polri Diaudit	Hukum
18	Viral Maling Cokelat di Alfamart, DPR: UU ITE Jangan Dipakai Menindas Orang Kecil	Hukum
19	Diduga Terima Suap, Rektor Unila Punya Harta Rp3,18 Miliar	Hukum
20	Baru Bebas dari Lapas Sukamiskin, Eks Wali Kota Cimahi Ajay Priatna Ditangkap KPK	Hukum
21	Tiket Pesawat Masih Mahal, Sandiaga Prediksi Tarif Turun Akhir Tahun	Travel
22	Promo Tiket Garuda Indonesia, Jakarta-Labuan Bajo PP Rp 1,7 Jutaan	Travel
23	HUT Ke-77 RI, Bendera Merah Putih Raksasa Dibentangkan di Gunung Bromo	Travel
24	5 Tempat Wisata yang Gratis Saat HUT Ke-77 RI	Travel
25	Syarat Perjalanan Terbaru, Vaksin Baru 2 Dosis Harus PCR	Travel
26	Kenali Apa itu Niacinamide, Kandungan, Fungsi, dan Efek Sampinya	Kesehatan
27	Apa Itu Batu Empedu?	Kesehatan
28	Punya Gizi yang Tinggi, Amankah Makan Telur Setiap Hari?	Kesehatan
29	4 Jenis Makanan untuk Turunkan Risiko Kanker Payudara	Kesehatan
30	Cara Mengatasi Infeksi Gigi Bungsu	Kesehatan
31	Top Gun Maverick Geser Capaian Box Office Titanic di Amerika	Hiburan
32	Review Manga: One Piece 1.056	Hiburan
33	Sinopsis King Arthur: Legend of the Sword di Bioskop Trans TV Hari Ini	Hiburan

34	Putus, Kim K Disebut Belum Siap Serius dengan Pete Davidson	Hiburan
35	FOTO: 'Konser' dari Para Kucing	Hiburan
36	Kalteng Siap Gelar Piala Dunia MTB 2022	Olahraga
37	Anthony Ginting di Kejuaraan Dunia 2022: Semoga Ketemu Axelsen	Olahraga
38	Perhatian Vladimir Putin pada Sambo	Olahraga
39	Momen Pengangkatan Atlet Berprestasi Jadi ASN	Olahraga
40	Putri KW Dipuji Jelang Kejuaraan Dunia 2022	Olahraga

Lampiran III

Daftar data artikel berita

Judul	Kategori
Pertamina Jaga Pasokan	Ekonomi
Dirut PP: Perseroan Beri Kesempatan Kerja dan Kesetaraan Karyawan Disabilitas	Ekonomi
In Picture: Pengrajin Sabun Berbahan Baku Minyak Jelantah	Ekonomi
Digitalisasi Dorong Daya Saing UMKM Lokal di Pasar Global	Ekonomi
Ekonom Sebut Pencairan Gaji ke-13 ASN Sangat tepat di Bulan Juli	Ekonomi
In Picture: Peternak Sugar Glider di Kabupaten Bogor	Ekonomi
Hadapi Idul Adha, Pemkot Yogyakarta Pastikan Ketersediaan Pangan	Ekonomi
IFC Berikan Pinjaman 300 Juta Dolar AS untuk Bank KB Bukopin	Ekonomi
Sri Mulyani: Alokasi Anggaran Pendidikan Tahun Depan Rp 608,3 Triliun	Ekonomi
Teten Yakin Harga Minyak Makan Merah Lebih Murah dari Minyak Goreng Biasa	Ekonomi
Harga Properti Terus Meningkat Meski Terbatas	Ekonomi
Harga Tanah Jadi Kendala Utama Masyarakat Bawah Miliki Rumah	Ekonomi
Dorong Kemajuan UMKM, Digital Entrepreneur Academy Resmi Diluncurkan	Ekonomi
Suharso Promosikan Peta Jalan Ekonomi Biru RI di Side Event G20	Ekonomi
Sri Mulyani Yakin AS dan Eropa Terjerumus ke Jurang Resesi Tahun Ini	Ekonomi
Anies Guyur Rp62,1 M Agar Tarif Transjakarta Tak Naik Usai BBM Mahal	Ekonomi
Harga Pangan Nasional Stabil Usai Harga BBM Naik Sepekan	Ekonomi
Malaysia Pastikan Tidak Ada Kandungan Alergen dalam Kecap dan Saus ABC	Ekonomi
KAI Percepat Waktu Tempuh 10 Perjalanan Kereta Jarak Jauh	Ekonomi
6 Daftar Bahan Pokok Penyumbang Terbesar Angka Kemiskinan di RI	Ekonomi
Jokowi Klaim 5,9 Juta Rakyat Miskin Sudah Terima BLT BBM	Ekonomi
Mensos Minta Pos Indonesia Salurkan BLT BBM ke 18,58 Juta Keluarga	Ekonomi

BPH Migas Sabet Penghargaan di Ajang ASEAN Energy Awards 2022	Ekonomi
Jokowi Bidik Ekonomi Tumbuh 5,3% di 2023, Pengamat Singgung Penurunan Belanja Negara	Ekonomi
Ungkap Capaian Positif Ekonomi dan Kekayaan Budaya, Menko Airlangga: Indonesia Kian Diperhitungkan Dunia	Ekonomi
Pacu Peran Generasi Muda untuk Percepat Pemulihan Ekonomi Daerah Pasca Pandemi	Ekonomi
Bantu Tingkatkan Ekonomi Masyarakat Pesisir, Indra Karya Raih 2 Penghargaan	Ekonomi
Pertumbuhan Ekonomi yang Impresif dan Kepemimpinan RI di Presidensi G20 jadi Perhatian Dunia	Ekonomi
Jokowi Dapat Bisikan Ekonomi Dunia Bakal Gelap, Wamenkeu: Tidak Terelakkan	Ekonomi
Ekonomi Indonesia Tumbuh Impresif Bukan Cuma Berkah Harga Komoditas	Ekonomi
Ekonomi Tumbuh 5,4%, Sri Mulyani: So, Inilah Bukti Proses Pemulihan	Ekonomi
Ekonomi RI Tumbuh 5,4%, Airlangga Minta BI Tahan Suku Bunga Acuan	Ekonomi
Airlangga Optimistis Ekonomi Indonesia Tumbuh di Atas 5% di Kuartal II 2022	Ekonomi
Ketakutan Terhadap Ekonomi China Bikin Rupiah Melemah	Ekonomi
Menguak 3 Alasan IMF Melukiskan Gambaran Suram Ekonomi Global di Sisa 2022	Ekonomi
Amerika Resesi, Ini 3 Dampaknya ke Ekonomi Indonesia	Ekonomi
Besok, Webinar Partai Perindo Bongkar Habis soal Moral Hazard Politik Uang di Level Voters	Politik
Berstatus Tahanan Kota, Habib Rizieq: Tak Kurangi Semangat Selamatkan Indonesia	Politik
Perempuan PAN Didorong Persiapkan Diri Hadapi Pemilu 2024	Politik
Akan Ada Parpol Gabung KIB, Golkar: Tidak Mengubah dan Mengganggu Irama Koalisi	Politik
Kader Partai Perindo Diminta Peka dan Peduli Persoalan Masyarakat	Politik
LaNyalla Sebut Kesetaraan Gender Masih di Atas Kertas, Perlu Kebijakan Spesifik	Politik
Alissa Wahid Minta Bebasnya Habib Rizieq Tak Dikaitkan dengan Kriminalisasi Ulama	Politik
Demokrat Sebut Surya Paloh Dukung AHY Maju Pilpres 2024, Tapi Jangan Ngoyo	Politik
Sandiaga Uno Siap Bangun Kerja Sama dengan Partai Perindo	Politik
Ratusan Milenial di Depok Dukung Sandiaga Uno Maju di Pilpres 2024	Politik
Jokowi dan Megawati Jadi Penentu di Pilpres 2024	Politik
Wujudkan Indonesia Emas 2045, Ribuan Mahasiswa Trisakti Dibekali Nilai-nilai Pancasila	Politik
Survei Indopol: Duet Prabowo-Muhaimin Unggul dalam Simulasi 3 Paslon	Politik

Survei Capres versi Indopol: Ganjar Dominasi Jateng-DIY, Anies Unggul di Jakarta	Politik
Kartu Prakerja Diubah Jadi Pelatihan Offline, PKS: Banyak Catatan dari KPK dan BPK	Politik
Kemendagri Kawal Pembentukan 3 Provinsi Baru di Papua	Politik
Ribuan Srikandi Lampung Dukung Ganjar lewat Cinta, Rasa, dan Warna Milenial	Politik
Survei Charta Politika: Elektabilitas Ganjar Pranowo Menang Telak di Jateng dan Jatim	Politik
Anak Muda Kaltim Pedes Ganjar Jadi Presiden 2024 dan Selesaikan Pembangunan IKN	Politik
Data Terbaru dari KPU, 38 Parpol Tingkat Nasional Miliki Akses Sipol	Politik
Survei Capres 2024: Ganjar, Prabowo, dan Anies Belum Goyah di Posisi Teratas	Politik
Bambang Pacul Yakin Ganjar Tak Akan Menang di Jateng Jika Diusung Partai Lain	Politik
DPR Minta Polisi Perhatikan Psikologis Anak dan Istri Ferdy Sambo	Politik
Mahfud MD: Kegagalan Pemerintah Banyak Ditiupkan di Medsos yang Sangat Brutal	Politik
Tak Berniat Nyalon Gubernur, Menteri Basuki: Sudah 70 Tahun	Politik
Partai Golkar Berkurban, Bagikan 93 Sapi dan 48 Kambing	Politik
Jokowi Resmikan Perluasan Bandara Komodo	Politik
Pilpres 2024, PPP Munculkan Pasangan Anies-Khofifah	Politik
Penjelasan Detail Lambang Sila ke-3 Pancasila	Politik
Letkol Inf Romel Jangga Wardhana Dilantik sebagai Dangrup 1 Kopassus	Politik
Penyaluran BLT BBM Tidak Merata, Partai Perindo: Perbaiki Data dan Perketat Pengawasan	Politik
Effendi Symbolon Dipanggil MKD Siang Ini, Pelapor Tiga Jam Lebih Dulu	Politik
Profil Effendi Symbolon, Politikus PDIP yang Sedang 'Diserbu' Prajurit TNI	Politik
Sekjen PPP Ungkap Suharso dan Mardiono Telah Bertemu Empat Mata, Ini yang Dibahas	Politik
Marak Hacker Bobol Data Pemerintahan, Faldo: Tak Ada Rahasia, Negara Ini Sudah Terbuka	Politik
BEM UI Sesalkan Pemerintah Naikkan Harga BBM, Kritik Penyaluran yang Salah Sasaran	Politik
Klaim Tidak Ada Perpecahan di PPP, Mardiono Sebut Suharso Monoarfa Sahabat	Politik
Marak Peretasan Data, Fadli Zon: Indonesia Dipermalukan Hacker	Politik
Cak Imin: Kader NU Jangan Wapres Terus, Sekali-kali Jadi Presiden	Politik
Kemenkumham Sahkan Mardiono Plt Ketua Umum DPP PPP	Politik
Pemuda Madiun Dipulangkan, Bjorka Sempat Sebut Pemerintah Salah Info	Teknologi

5 Penyebab Utama Power Steering Rusak	Teknologi
Angkatan Laut AS Ogah Buka Video-video UFO buat Publik, Kenapa?	Teknologi
Gaikindo: Baterai Buatan Indonesia Siap Produksi 2024	Teknologi
Bjorka Ejek Dark Tracer Beri Informasi Palsu ke Pemerintah	Teknologi
Deret Pejabat RI Kena 'Colek' Bjorka di Medsos: Plate Hingga Anies	Teknologi
Honda Produksi 5 Juta Transmisi MCVT dalam 9 Tahun, Sudah Diekspor?	Teknologi
NASA Kebut Perbaiki Artemis 1 Jelang Peluncuran Ulang Akhir Pekan Ini	Teknologi
Asosiasi Operator soal 1 NIK Dipakai Ribuan SIM Card: Saya Tak Tahu	Teknologi
Data Agen Intel BIN Diduga Bocor Lagi, RI Disindir Negara Open Source	Teknologi
AS Larang Ekspor Semikonduktor Canggih pada Teknologi Manufaktur	Teknologi
Gara-gara Ini! Pendiri Telegram Protes Keras ke Apple	Teknologi
Cara Jualan di TikTok Shop, Dijamin Produk Anda Laris Manis!	Teknologi
Startup Luar Negeri Rasa Indonesia, Techbros Perkenalkan Teknologi TechbrosFi Mobile Private Network	Teknologi
HP dengan Snapdragon Harga Rp2 Jutaan, Ini yang Terbaik!	Teknologi
Gunakan Sistem Baru, LYTO Luncukan Game Dekaron MMORPG	Teknologi
Cara Mencari Masjid Terdekat Menggunakan Google Maps	Teknologi
Google Sempat Down Selama 1 Jam, Begini Dampaknya	Teknologi
5 Fakta Google Maps yang Belum Banyak Diketahui Orang	Teknologi
Samsung Akan Gunakan Teknologi UI Terbaru untuk Galaxy S22	Teknologi
Serangan Siber ke Perusahaan Gaming Melebihi Sektor Finansial	Teknologi
Meta Hapus Ribuan Akun Palsu Pro-Rusia di Facebook dan Instagram	Teknologi
Rekomendasi Laptop 2 in 1 Termurah 2022, Beli Laptop Bonus Tablet!	Teknologi
Tips Merawat Baterai Laptop Agar Lebih Awet	Teknologi
Alasan LG PuriCare Wearable Air Purifier Cocok Digunakan Tenaga Medis	Teknologi
Perbandingan Harga dan Spesifikasi Samsung S22 Ultra dan Samsung S22 Plus	Teknologi
Awalnya Anggap Media Sosial Adalah Lelucon, Seniman Grafiti Ini Malah Jadi Kaya dari Facebook	Teknologi
Cegah Gangguan Internet, APJII Dan Matrix NAP Berkolaborasi	Teknologi
Xiaomi Kenalkan Ikat Kepala Canggih, Bisa Kontrol Perangkat Pintar Lewat Pikiran	Teknologi
Ini Keunikan Desain dan Kamera OPPO Reno8	Teknologi
Kominfo Putus Akses 15 PSE Game Judi Online	Teknologi
Mengapa Judi Online Tidak Diblokir Kominfo?	Teknologi
Kominfo Minta Bantuan Kedutaan Amerika Agar PayPal, Steam, Dota, dan CS: GO Mau Daftar PSE	Teknologi
Tak Berani Tindak Tegas Google dan WhastApp Cs, #BlokirKominfo Kuasai Twitter	Teknologi

Spesifikasi dan Harga HP Nothing Phone di Indonesia	Teknologi
Kominfo: Layanan PayPal Dibuka Sementara Hari Ini hingga Jumat, 5 Agustus 2022	Teknologi
Fitur Layanan Keuangan Digital BNC Terus Bertransformasi	Teknologi
Developer Game Stray Bekerjasama untuk Rawat Kucing Jalanan	Teknologi
Bergerak ke Arah Seni dan Teknologi, NFT for Everyone Resmi Digelar	Teknologi
YouTube Kini Bisa Ubah Konten Durasi Panjang Jadi Shorts	Teknologi
5 Baju Adat Jokowi Saat Pidato Kenegaraan Sidang Tahunan MPR RI	Travel
Panduan Wisata Gembira Loka Zoo Yogyakarta: Harga Tiket dan Jam Buka	Travel
Bentang Alam Raja Ampat yang Rupawan	Travel
Festival Golo Koe Tampilkan The New Labuan Bajo kepada Dunia	Travel
Visa Turis Bisa untuk Umrah, Tak Berlaku bagi Jemaah Indonesia	Travel
Tiket Pesawat Masih Mahal, Sandiaga Prediksi Tarif Turun Akhir Tahun	Travel
Promo Tiket Garuda Indonesia, Jakarta-Labuan Bajo PP Rp 1,7 Jutaan	Travel
HUT Ke-77 RI, Bendera Merah Putih Raksasa Dibentangkan di Gunung Bromo	Travel
5 Tempat Wisata yang Gratis Saat HUT Ke-77 RI	Travel
Syarat Perjalanan Terbaru, Vaksin Baru 2 Dosis Harus PCR	Travel
Makam Bung Karno di Blitar, Tempat Peristirahatan Terakhir Sang Proklamator	Travel
Kereta Zombie di LRT Jakarta Bakal Tambah Konsep Baru, Seperti Apa?	Travel
6 Tips Main ke Wahana Train to Apocalypse di Kelapa Gading	Travel
Pasar Seni Desa Kenalan Borobudur, Satukan Air Suci dari 3 Sumber Berbeda	Travel
Jelang HUT Ke-77 RI, Kuota Pendakian Gunung Binaiya Maluku Penuh	Travel
20.000 Peserta Ditargetkan Ikut Upacara HUT Ke-77 RI di Gunung Talang	Travel
Pesona Danau Pading, Bekas Tambang yang Jadi Wisata di Babel	Travel
Pengalaman Nikmati Kuliner di Marawa Beach Club, Wisata Baru di Padang	Travel
Tempat Wisata Pecinan Kya-kya di Kota Surabaya Akan Dibuka	Travel
Lumba-lumba Muncul di Pantai Tiga Warna Malang, Dianggap Kejadian Langka	Travel
5,9 Juta Wisatawan Kunjungi Malang Sepanjang 2022	Travel
15 Hotel Dekat Malioboro, Instagramable dan Nyaman buat Nginap	Travel
Berwisata ke Bantul Saat Bulan Suro, Ada Banyak Event Budaya	Travel
Kemenparekraf Siapkan Pola Perjalanan untuk Wisatawan ke TN Komodo	Travel
Daftar Kalender Kegiatan Pariwisata Sumatera Barat Agustus-Akhir 2022	Travel
Kenaikan Tiket Taman Nasional Komodo Ditunda, Catat Harganya Saat Ini	Travel
7 Wisata Salatiga, Ada Tempat Liburan dengan Pemandangan Alam	Travel

Syarat Naik Kano Gratis di Kali Sipon Tangerang, Siap Antre	Travel
Terupa Festival di Bali, Tampilkan Puluhan Karya Seni NFT	Travel
Balikipapan akan Punya Wisata Susur Teluk dengan Kapal Pinisi	Travel
Masuk Thailand Kini Tak Perlu Tes PCR dan Syarat Lainnya	Travel
Gusti Mangkubumi dan Galeri Keris di Dalem Poenakawan	Travel
Jakarta-Bali, Salah Satu Rute Penerbangan Domestik Tersibuk di Dunia	Travel
Potensi Pariwisata di Malang Barat, Wisata Kebun hingga Rafting	Travel
Menara Eiffel di Rawa Pening Roboh Diterjang Badai	Travel
Kerja Sama Indonesia-Belanda Diharapkan Bisa Majukan Wisata Sejarah	Travel
Batik Air Tambah Frekuensi Penerbangan Jakarta-Singapura PP	Travel
Desa Wisata Pentagen Jambi, Punya Taman Wisata Air	Travel
Pagoda Pantjoran PIK Buka untuk Umum, Bisa Masuk Lantai 1	Travel
Prof. Martono Terpilih Jadi Rektor Unnes Periode 2022-2026	Travel
UT Perkuat Literasi Digital Guru Tangsel lewat Pengabdian Masyarakat	Pendidikan
Mahasiswa UAJY Ciptakan Salep Daun Jamblang untuk Atasi Jerawat	Pendidikan
UI Satu-satunya Kampus Indonesia Penerima Penghargaan QS Recognition for Engagement	Pendidikan
Indonesia Jadi Tuan Rumah Olimpiade Informatika Internasional 2022	Pendidikan
Cerita Maizidah Salas, Dirikan PAUD Inklusi bagi Anak TKI di Wonosobo	Pendidikan
Sempat Diremehkan, Kini Siswa SMK Ini Buktikan Kuliah di ITB	Pendidikan
Guru Besar IPB: Ini Manfaat Kedelai untuk Kesehatan	Pendidikan
Dukung Pendidikan dan SDM Indonesia, Vista Education Gelar Pameran Worldwide Education Fair 2022	Pendidikan
Tim ITS Kalahkan Kampus Nomor 1 Dunia di Ajang Ini	Pendidikan
3 Mahasiswa UM Surabaya Raih 7 Medali ASEAN University Games	Pendidikan
Olimpiade Sains Nasional 2022 Jenjang SD dan SMP Resmi Dibuka	Pendidikan
Kurikulum Merdeka dan Perlunya Persiapan Sekolah	Pendidikan
IPB Mulai Jaring Calon Rektor Periode 2023-2028	Pendidikan
Cerita Mahasiswi yang Tertarik Dunia Robotik	Pendidikan
4 Fakta tentang Vaksin Cacar Monyet yang Disetujui FDA	Pendidikan
Beasiswa Aktivist Salman Dibuka untuk Mahasiswa D3/S1, Ini Syaratnya	Pendidikan
PTS Minta Pembebasan Pajak untuk Sarana Laboratorium	Pendidikan
Unpad Ungkap 3 Alasan Dukung Perubahan Skema Penerimaan Mahasiswa Baru	Pendidikan
PTS Usul Perguruan Tinggi Asing di Indonesia Dibatasi, Ada Berapa Sekarang?	Pendidikan
Dapat Beasiswa Program IISMA, 10 Mahasiswa Indonesia Siap Kuliah di Belgia	Pendidikan

Siswa Bisa Pilih Prodi Lintas Jurusan di Seleksi Masuk PTN 2023, tapi...	Pendidikan
7 Zat Kimia yang Bersifat Korosif	Pendidikan
20 SMA Terbaik di Jakarta Timur Versi Nilai UTBK 2022, Ada Sekolahmu?	Pendidikan
4 Keunggulan Politeknik Dibandingkan dengan Universitas, Kamu Tertarik?	Pendidikan
Telkom University Buka Jalur Beasiswa Anak Guru 2023, Cek Syaratnya di Sini	Pendidikan
17 PTKIN Terbaik Versi UniRank 2022, UIN Sunan Kalijaga Juara 1	Pendidikan
Peneliti ITB: Belajar sambil Bermain Paling Cocok untuk Mendidik Generasi Z	Pendidikan
Tips Raih Akreditasi Unggul ala Mantan Mendikbud Prof. Bambang Sudibyo	Pendidikan
Berusia 15 Tahun, Raja Dinobatkan Jadi Mahasiswa Termuda UGM	Pendidikan
Universitas Bakrie: Masa Depan Lebih Terjamin dengan Pendidikan Tinggi Berkualitas	Pendidikan
UGM dan Politeknik Pertanian Siap Lahirkan Petani Modern dengan Inovasi Terbaru	Pendidikan
Galang Donasi Beasiswa, Unika Atma Jaya Gelar Lomba Lari	Pendidikan
Dorong Ekonomi Sirkular, Mahasiswa ITS Gagas Waralaba Pengolahan Kompos	Pendidikan
Cerita Inspiratif Salsabilla, Wisudawan Teman Tuli yang Sukses Lulus dari ITB	Pendidikan
Viral Curhatan Wanita yang Berjuang Dapatkan 4 Gelar Sekaligus dari UI, UGM, BINUS, dan USYD	Pendidikan
UI, ASEAN University Network, dan PPI Dunia Gelar Webinar Penanganan Covid-19	Pendidikan
7.539 Peserta Seleksi Mandiri IPB University Ikuti Ujian Secara Online	Pendidikan
Cerita Mahasiswa ITS Studi di Korsel dan Terkesan dengan Budaya Tepat Waktu	Pendidikan
Mahasiswa Unair Ciptakan Bahan Bakar Alternatif dari Sekam	Pendidikan
Kisah Sukses Alumni UPER: Raih Beasiswa hingga Kerja di Perusahaan Ternama	Pendidikan
Siap Tur ke AS, D'Masiv Punya Strategi Gaet Penonton Bule	Hiburan
Arctic Monkeys Kembali Naik Panggung Usai 3 Tahun	Hiburan
Sinopsis 12 Strong Tayang di Bioskop Trans TV Malam Ini	Hiburan
Viral Pria Mirip Johnny Depp Nimbrung di Hari Karbala Iran	Hiburan
BLACKPINK Akan Konser Born Pink di Jakarta 11 Maret 2023	Hiburan
Top Gun Maverick Geser Capaian Box Office Titanic di Amerika	Hiburan
Review Manga: One Piece 1.056	Hiburan
Sinopsis King Arthur: Legend of the Sword di Bioskop Trans TV Hari Ini	Hiburan
Putus, Kim K Disebut Belum Siap Serious dengan Pete Davidson	Hiburan
FOTO: 'Konser' dari Para Kucing	Hiburan

FOTO: Pentas Tari Tradisional, Merdeka Indonesiaku	Hiburan
Usaha Fan Meriahkan Konser 10 Tahun JKT48, Bagi-bagi Lighstick Gratis	Hiburan
5 Rekomendasi Film Akhir Pekan, Pengabdian Setan 2 Communion	Hiburan
Sinopsis KinnPorsche the Series, Drama Thailand Terpopuler di iQIYI	Hiburan
Kang Tae-oh, Cowok Idaman Bintang Extraordinary Attorney Woo	Hiburan
Cha Eun-woo ASTRO Diincar Perankan Pengacara dalam Drama Baru	Hiburan
Sutradara Respons Batgirl Batal Tayang: Masih Tidak Percaya	Hiburan
Joker 2 Akan Tayang 4 Oktober 2024	Hiburan
Tim Produksi Mulai Tulis Naskah Stranger Things 5	Hiburan
Justin Bieber Bahagia Konser Lagi Usai Alami Lumpuh Wajah	Hiburan
Film Live-Action Tokyo Revengers 2 Akan Tayang 2023	Hiburan
Sinopsis Film The Scout, Petaka Berburu Harta Karun di Kuil Tua	Hiburan
9 Rekomendasi Film Terbaru Tayang Agustus 2022	Hiburan
Everything Everywhere All at Once Cetak Rekor untuk A24 Secara Global	Hiburan
Ada Cerita Nyata di Balik Serial Yang Hilang Dalam Cinta	Hiburan
Seol In-ah Kepincut Kuliner Indonesia, Paling Suka Mie Goreng	Hiburan
Sinopsis Today's Webtoon, Mantan Atlet Jadi Editor Webtoon	Hiburan
FOTO: Wayang, Kepala Kerbau dan Tradisi Satu Suro Tahun Baru Jawa	Hiburan
Review Film: DC League of Super Pets	Hiburan
5 Rekomendasi Film Akhir Pekan, Alienoid	Hiburan
Sinopsis Spy x Family Part 2, Keselamatan Keluarga Forger Terancam	Hiburan
Daftar Lengkap Pemenang AMI Awards 2022	Hiburan
7 Film Horor Boneka, Terbaru The Doll 3 dari Indonesia	Hiburan
Review Film: Smile	Hiburan
Travis Japan Bicara Soal Branding hingga Nama Basis Penggemar	Hiburan
Sinopsis Assassin's Creed, Bioskop Trans TV 14 Oktober 2022	Hiburan
5 Rekomendasi Film Akhir Pekan, Don't Worry Darling hingga Inang	Hiburan
SEVENTEEN Akan Konser Lagi di Indonesia 28 Desember	Hiburan
Sinopsis Halloween Ends, Akhir Pertarungan Laurie vs Michael	Hiburan
Beyonce Bantah Tuduhan Tak Bayar Desainer untuk Album Renaissance	Hiburan
Warisan Shinzo Abe di Olahraga: Olimpiade Tokyo 2020	Olahraga
Hasil Malaysia Masters: Gregoria Singkirkan Nomor Satu Dunia Akane	Olahraga
Marselino Cerita Kronologi Cedera: Kaki Bunyi saat Paksa Passing	Olahraga
Stadion Piala Dunia 2022 Bebas Alkohol, Penjualan Bir di Luar	Olahraga
Juventus Pamer Kedatangan Di Maria di Turin	Olahraga
Petenis 'Buangan Rusia' Rybakina Juara Wimbledon 2022	Olahraga

Chico Juara Malaysia Masters, Hapus Dahaga Gelar Tunggal Putra	Olahraga
Hasil UFC: Said Nurmagomedov Menang Angka Mutlak atas De Andrade	Olahraga
4 Wakil Indonesia di Final Malaysia Masters, 1 Gelar di Tangan	Olahraga
Chico Sempat Salah Strategi dalam Perjuangan ke Final Malaysia Masters	Olahraga
Timnas Indonesia Lebih Buas di Babak Pertama di Piala AFF U-19	Olahraga
EDUSPORTS: Sejarah Bendera Asisten Wasit Sepak Bola	Olahraga
Taktik Tak Biasa dan Standar Tinggi STY di Timnas Indonesia U-19	Olahraga
Rahasia Gregoria Dua Kali Beruntun Tekuk Akane	Olahraga
Lampu Stadion Patriot Sempat Mati di Thailand vs Brunei	Olahraga
MotoGP Austria 2022: Menantang Bagnaia Kembalikan Dominasi Ducati	Olahraga
Kejuaraan Dunia Bulutangkis 2022: Gregoria Vs Tekanan Mental	Olahraga
Merdeka dalam Olahraga: Berprestasi, Menyemangati, Mengapresiasi	Olahraga
Conte Mencak-mencak Lawan Chelsea, Pemain Tottenham Sudah Biasa	Olahraga
Mental Putri KW Sudah Terasah Tatap Kejuaraan Dunia Bulutangkis 2022	Olahraga
Kalteng Siap Gelar Piala Dunia MTB 2022	Olahraga
Anthony Ginting di Kejuaraan Dunia 2022: Semoga Ketemu Axelsen	Olahraga
Perhatian Vladimir Putin pada Sambo	Olahraga
Momen Pengangkatan Atlet Berprestasi Jadi ASN	Olahraga
Putri KW Dipuji Jelang Kejuaraan Dunia 2022	Olahraga
Petarung UFC Wanita Ini Lelang Celana Dalamnya demi...	Olahraga
Johann Zarco Berniat Segera Move On dari Kegagalan di MotoGP Inggris	Olahraga
Duel Sengit Bagnaia vs Vinales di Dua Lap Terakhir	Olahraga
Klasemen MotoGP 2022: Bagnaia Dekati Quartararo dan Aleix Espargaro	Olahraga
GFNY 2022 Kembali Bergulir di Bali	Olahraga
MU Belanja Besar, Ten Hag Rasakan Tekanan Enggak?	Olahraga
Madrid Vs Betis: El Real Masih Sempurna Usai Menang Tipis 2-1	Olahraga
Persija Vs Bhayangkara FC: Krmencik Dua Gol, Macan Kemayoran Menang 2-1	Olahraga
Tottenham Sudah Belanja Rp 2,9 Triliun, tapi Conte Belum Puas	Olahraga
Agen Ruediger Tepis Tudingan Kliennya Serakah	Olahraga
Marko Simic Akhirnya Dapat Klub Baru Usai Tinggalkan Persija	Olahraga
Chelsea Dekati Manajer Brighton Graham Potter	Olahraga
Liga Champions Keras, Bos! Dua Pelatih Dipecat Usai Laga Pertama	Olahraga
Ronaldo dan Maguire Jadi Starter di Liga Europa?	Olahraga
Resmi! MU Gaet Martin Dubravka	Olahraga
Bagaimana Cara Pasien Cacar Monyet Bisa Sembuh?	Kesehatan
Cacar Monyet Masuk Indonesia, IDI Minta Masyarakat Tidak Panik	Kesehatan

Teh Tidak Bisa Sepenuhnya Berguna untuk Menurunkan Berat Badan	Kesehatan
Sering Disepelekan, Begini Bahaya Flu	Kesehatan
Kenali Apa itu Niacinamide, Kandungan, Fungsi, dan Efek Sampingnya	Kesehatan
Apa Itu Batu Empedu?	Kesehatan
Punya Gizi yang Tinggi, Amankah Makan Telur Setiap Hari?	Kesehatan
4 Jenis Makanan untuk Turunkan Risiko Kanker Payudara	Kesehatan
Cara Mengatasi Infeksi Gigi Bungsu	Kesehatan
Penyakit Polip Rahim, Bahaya atau Tidak?	Kesehatan
Fungsi Paru-paru dan Cara Menjaga Kesehatannya	Kesehatan
4 Makanan untuk Penderita Polip Rahim dan Pantangannya	Kesehatan
Apakah Orang Dapat Mencegah Risiko Stroke karena Keturunan?	Kesehatan
Fungsi Ginjal dan Cara Menjaga Kesehatannya	Kesehatan
5 Tips Hubungan Seks yang Aman Setelah Operasi Caesar	Kesehatan
10 Gejala Paru-paru Bocor yang Pantang Disepelekan	Kesehatan
Amankah Konsumsi Suplemen Antioksidan untuk Penderita Kanker?	Kesehatan
Kenali Apa itu Infeksi HPV, Penyebab Kanker Serviks dan Kutil Kelamin	Kesehatan
Apakah TBC Otak Bisa Menular?	Kesehatan
Alat Deteksi Dini Kanker Serviks Non Invansif Segera Diluncurkan	Kesehatan
Apakah Penyakit Pankreatitis Berbahaya?	Kesehatan
Kenali Apa itu Strict Parents, Tanda, dan Dampaknya pada Anak	Kesehatan
Gejala Hepatitis Alkoholik yang Harus Diwaspadai	Kesehatan
Apa Fungsi Jantung?	Kesehatan
8 Cara Alami Turunkan Berat Badan: Konsumsi Protein hingga Tidur Cukup	Kesehatan
Manfaat Ganja Medis dan Efek Sampingnya bagi Kesehatan	Kesehatan
Mengenal Biopsi Payudara, Tahapan Pemeriksaan Kanker	Kesehatan
Bagaimana Kanker Payudara Memengaruhi Kesehatan Tubuh?	Kesehatan
Apakah Hipertiroid Bisa Sembuh?	Kesehatan
Penyakit Hepatitis B Menular Lewat Apa Saja?	Kesehatan
Stres Dapat Memicu Kerusakan Ginjal, Kenali Penyebab dan Pencegahannya	Kesehatan
Cara Kerja Statin untuk Menurunkan Kolesterol	Kesehatan
Manfaat dan Efek Samping Suplemen Diet untuk Kesehatan	Kesehatan
7 Makanan untuk Membantu Mengatasi Anemia karena Defisiensi Zat Besi	Kesehatan
Tanda-tanda Kanker Ginjal yang Harus Diwaspadai	Kesehatan
Kenali Stadium Kanker Ginjal untuk Mengetahui Risikonya	Kesehatan
5 Obat Alami untuk Mengatasi Berbagai Penyakit Kulit	Kesehatan
Kenali Apa itu Direct Breastfeeding pada Bayi, Manfaat, dan Caranya	Kesehatan

Komplikasi Penyakit Asam Urat yang Tidak Bisa Dianggap Enteng	Kesehatan
OTT Rektor Unila Terkait Dugaan Suap Penerimaan Mahasiswa Baru Jalur Mandiri	Hukum
Komnas HAM Mulai Susun Kerangka Peristiwa Kasus Kematian Brigadir J	Hukum
Belum Laporan ke KPK Soal Amplop Cokelat dari Ferdy Sambo, Begini Penjelasan LPSK	Hukum
Komnas Perempuan dan Komnas HAM Dorong Pendampingan Psikologi Putri Candrawathi	Hukum
Kasat Narkoba Polres Karawang Pernah Antar 2.000 Butir Ekstasi ke Tempat Hiburan Malam	Hukum
KPK Tetapkan Mantan Kepala BPKAD Jatim Tersangka Korupsi Bantuan Keuangan	Hukum
Komisi III DPR Minta Satgassus Polri Diaudit	Hukum
Viral Maling Cokelat di Alfamart, DPR: UU ITE Jangan Dipakai Menindas Orang Kecil	Hukum
Diduga Terima Suap, Rektor Unila Punya Harta Rp3,18 Miliar	Hukum
Baru Bebas dari Lapas Sukamiskin, Eks Wali Kota Cimahi Ajay Priatna Ditangkap KPK	Hukum
Surya Darmadi Penuhi Panggilan Kejagung, Ini Respons KPK	Hukum
KPK Periksa Sejumlah Pejabat Pemalang terkait Kasus Jual Beli Jabatan	Hukum
HUT ke-77 RI, Kemenkumham: 168.196 Narapidana Terima Remisi, 2.725 Langsung Bebas	Hukum
KPK Amankan Uang Tunai Usai Geledah Kantor hingga Rumah Pribadi Bupati Pemalang	Hukum
Komnas HAM Yakini Ada Dugaan Kuat Pelanggaran HAM dalam Kasus Tewasnya Brigadir J	Hukum
HDKD ke-77, DPR Harap Kemenkumham Jadi Mitra Kerja Kolaboratif	Hukum
DPR Apresiasi Perjuangan Kepala BP2MI Lindungi Pekerja Migran Indonesia	Hukum
Surya Darmadi Pulang ke Indonesia 15 Agustus, SA Institut: Bagian dari Prestasi Kejaksanaan	Hukum
Ketua MPR RI Luncurkan 2 Buku Baru, HT: 4 Pilar Harus Disosialisasikan	Hukum
Kemhub Izinkan Harga Tiket Pesawat Naik, Perindo: Kebijakan yang Kurang Bijak	Hukum
Resmi Mendaftar di KPU, HT: Perindo untuk Indonesia Sejahtera	Hukum
UU Tentang Sumbar Dinilai Berpotensi Melanggar Hak Minoritas	Hukum
Rapimnas V KAHMI Hasilkan Sejumlah Pokok Pikiran	Hukum
Perludem Sebut Koalisi Politik Sejak Dini Untungkan Rakyat dan Partai	Hukum
Wakil Ketua DPR Malaysia Apresiasi Kunjungan Ketum KNPI	Hukum
PP Muhammadiyah Dorong Pemerintah-DPR Bikin Regulasi Pengawasan Lembaga Filantropi	Hukum

Cak Imin di Mata Pegiat HAM: Representasikan Sikap Moderat NU, Mampu Jaga Kebhinekaan	Hukum
Soal RUU KUHP, Begini Catatan Kritis Guru Besar Hukum Pidana UI Indriyanto Seno Adji	Hukum
Wakil Ketua MPR Tekankan Pentingnya UU PPRT untuk Lindungi Pekerja Migran Indonesia	Hukum
Peningkatan Capaian Polri, DPR Sebut Presisi Bukan Sekadar Slogan	Hukum
Mantan Bendahara Demokrat Balikpapan Dieksekusi ke Lapas Perempuan Tenggara	Hukum
Tiba di Bandara Soetta, 3 DPO Kasus Judi Online Dibawa ke Bareskrim	Hukum
Penangkapan Teddy Minahasa, Komisi III DPR: Polri Jangan Pandang Bulu	Hukum
Jelang Kedatangan Apin BK, Pasukan Polisi Bersenjata Lengkap Tiba di Bandara Soetta	Hukum
Bandar Judi Kakap Apin BK Dijadwalkan Tiba di Bandara Soetta Malam Ini	Hukum
Kasus Narkoba Teddy Minahasa, Polda Metro Tangkap Polisi Berpangkat Briptu, Kompol, dan AKBP	Hukum
Sampaikan Laporan ke Presiden, TGIPF Sebut Banyak Temuan Indikasi Dalam Tragedi Kanjuruhan	Hukum
TGIPF: Presiden Jokowi Minta Polri Lanjutkan Penyelidikan Tragedi Kanjuruhan	Hukum
Kapolda Jatim Ditangkap Propam Kasus Narkoba, Komisi III DPR: Lebih Cepat Ketahuan Lebih Baik	Hukum
Komnas HAM Ngaku Kantongi Bukti Video Eksklusif Pintu 13 Stadion Kanjuruhan Terbuka	Hukum

Lampiran IV

Hasil klasifikasi pengujian 90:10

No.	Ground truth	Prediksi	TP	TN	FP	FN
1	Politik	Politik	1	8	0	0
2	Politik	Politik	1	8	0	0
3	Politik	Politik	1	8	0	0
4	Politik	Politik	1	8	0	0
5	Teknologi	Teknologi	1	8	0	0
6	Teknologi	Teknologi	1	8	0	0
7	Teknologi	Hukum	0	7	1	1
8	Teknologi	Hukum	0	7	1	1
9	Travel	Travel	1	8	0	0
10	Travel	Travel	1	8	0	0
11	Travel	Travel	1	8	0	0
12	Travel	Travel	1	8	0	0
13	Hiburan	Hiburan	1	8	0	0
14	Hiburan	Hiburan	1	8	0	0

15	Hiburan	Hiburan	1	8	0	0
16	Hiburan	Olahraga	0	7	1	1
17	Olahraga	Olahraga	1	8	0	0
18	Olahraga	Olahraga	1	8	0	0
19	Olahraga	Olahraga	1	8	0	0
20	Olahraga	Teknologi	0	7	1	1
21	Pendidikan	Olahraga	0	7	1	1
22	Pendidikan	Pendidikan	1	8	0	0
23	Pendidikan	Pendidikan	1	8	0	0
24	Pendidikan	Pendidikan	1	8	0	0
25	Ekonomi	Ekonomi	1	8	0	0
26	Ekonomi	Ekonomi	1	8	0	0
27	Ekonomi	Ekonomi	1	8	0	0
28	Ekonomi	Kesehatan	0	7	1	1
29	Hukum	Hukum	1	8	0	0
30	Hukum	Hukum	1	8	0	0
31	Hukum	Hukum	1	8	0	0
32	Hukum	Hukum	1	8	0	0
33	Kesehatan	Kesehatan	1	8	0	0
34	Kesehatan	Kesehatan	1	8	0	0
35	Kesehatan	Kesehatan	1	8	0	0
36	Kesehatan	Kesehatan	1	8	0	0
Total			30	282	6	6

Lampiran V

Hasil klasifikasi pengujian 80:20

No.	Ground truth	Prediksi	TP	TN	FP	FN
1	Politik	Politik	1	8	0	0
2	Politik	Politik	1	8	0	0
3	Politik	Politik	1	8	0	0
4	Politik	Politik	1	8	0	0
5	Politik	Politik	1	8	0	0
6	Politik	Politik	1	8	0	0
7	Politik	Pendidikan	0	7	1	1
8	Politik	Pendidikan	0	7	1	1
9	Teknologi	Teknologi	1	8	0	0
10	Teknologi	Teknologi	1	8	0	0
11	Teknologi	Teknologi	1	8	0	0
12	Teknologi	Teknologi	1	8	0	0
13	Teknologi	Teknologi	1	8	0	0

14	Teknologi	Olahraga	0	7	1	1
15	Teknologi	Olahraga	0	7	1	1
16	Teknologi	Hukum	0	7	1	1
17	Travel	Travel	1	8	0	0
18	Travel	Travel	1	8	0	0
19	Travel	Travel	1	8	0	0
20	Travel	Travel	1	8	0	0
21	Travel	Travel	1	8	0	0
22	Travel	Travel	1	8	0	0
23	Travel	Travel	1	8	0	0
24	Travel	Travel	1	8	0	0
25	Hiburan	Hiburan	1	8	0	0
26	Hiburan	Olahraga	0	7	1	1
27	Hiburan	Olahraga	0	7	1	1
28	Hiburan	Olahraga	0	7	1	1
29	Hiburan	Hiburan	1	8	0	0
30	Hiburan	Hiburan	1	8	0	0
31	Hiburan	Hiburan	1	8	0	0
32	Hiburan	Hukum	0	7	1	1
33	Olahraga	Olahraga	1	8	0	0
34	Olahraga	Olahraga	1	8	0	0
35	Olahraga	Olahraga	1	8	0	0
36	Olahraga	Olahraga	1	8	0	0
37	Olahraga	Olahraga	1	8	0	0
38	Olahraga	Pendidikan	0	7	1	1
39	Olahraga	Pendidikan	0	7	1	1
40	Olahraga	Pendidikan	0	7	1	1
41	Pendidikan	Hukum	0	7	1	1
42	Pendidikan	Hukum	0	7	1	1
43	Pendidikan	Pendidikan	1	8	0	0
44	Pendidikan	Pendidikan	1	8	0	0
45	Pendidikan	Pendidikan	1	8	0	0
46	Pendidikan	Pendidikan	1	8	0	0
47	Pendidikan	Pendidikan	1	8	0	0
48	Pendidikan	Pendidikan	1	8	0	0
49	Ekonomi	Ekonomi	1	8	0	0
50	Ekonomi	Ekonomi	1	8	0	0
51	Ekonomi	Ekonomi	1	8	0	0
52	Ekonomi	Ekonomi	1	8	0	0
53	Ekonomi	Ekonomi	1	8	0	0

54	Ekonomi	Hukum	0	7	1	1
55	Ekonomi	Kesehatan	0	7	1	1
56	Ekonomi	Kesehatan	0	7	1	1
57	Hukum	Hukum	1	8	0	0
58	Hukum	Hukum	1	8	0	0
59	Hukum	Hukum	1	8	0	0
60	Hukum	Hukum	1	8	0	0
61	Hukum	Pendidikan	0	7	1	1
62	Hukum	Pendidikan	0	7	1	1
63	Hukum	Pendidikan	0	7	1	1
64	Hukum	Pendidikan	0	7	1	1
65	Kesehatan	Kesehatan	1	8	0	0
66	Kesehatan	Kesehatan	1	8	0	0
67	Kesehatan	Kesehatan	1	8	0	0
68	Kesehatan	Kesehatan	1	8	0	0
69	Kesehatan	Kesehatan	1	8	0	0
70	Kesehatan	Hukum	0	7	1	1
71	Kesehatan	Hukum	0	7	1	1
72	Kesehatan	Hukum	0	7	1	1
Total			48	552	24	24

Lampiran VI

Hasil klasifikasi pengujian 70:30

No.	Ground truth	Prediksi	TP	TN	FP	FN
1	Politik	Politik	1	8	0	0
2	Politik	Politik	1	8	0	0
3	Politik	Politik	1	8	0	0
4	Politik	Politik	1	8	0	0
5	Politik	Politik	1	8	0	0
6	Politik	Politik	1	8	0	0
7	Politik	Pendidikan	0	7	1	1
8	Politik	Pendidikan	0	7	1	1
9	Politik	Pendidikan	0	7	1	1
10	Politik	Pendidikan	0	7	1	1
11	Politik	Pendidikan	0	7	1	1
12	Politik	Pendidikan	0	7	1	1
13	Teknologi	Teknologi	1	8	0	0
14	Teknologi	Teknologi	1	8	0	0
15	Teknologi	Teknologi	1	8	0	0

16	Teknologi	Teknologi	1	8	0	0
17	Teknologi	Teknologi	1	8	0	0
18	Teknologi	Hiburan	0	7	1	1
19	Teknologi	Hiburan	0	7	1	1
20	Teknologi	Hiburan	0	7	1	1
21	Teknologi	Hukum	0	7	1	1
22	Teknologi	Hukum	0	7	1	1
23	Teknologi	Hukum	0	7	1	1
24	Teknologi	Hukum	0	7	1	1
25	Travel	Travel	1	8	0	0
26	Travel	Travel	1	8	0	0
27	Travel	Travel	1	8	0	0
28	Travel	Travel	1	8	0	0
29	Travel	Travel	1	8	0	0
30	Travel	Travel	1	8	0	0
31	Travel	Ekonomi	0	7	1	1
32	Travel	Ekonomi	0	7	1	1
33	Travel	Ekonomi	0	7	1	1
34	Travel	Ekonomi	0	7	1	1
35	Travel	Ekonomi	0	7	1	1
36	Travel	Ekonomi	0	7	1	1
37	Hiburan	Hiburan	1	8	0	0
38	Hiburan	Hiburan	1	8	0	0
39	Hiburan	Hiburan	1	8	0	0
40	Hiburan	Hiburan	1	8	0	0
41	Hiburan	Hiburan	1	8	0	0
42	Hiburan	Hiburan	1	8	0	0
43	Hiburan	Hiburan	1	8	0	0
44	Hiburan	Olahraga	0	7	1	1
45	Hiburan	Olahraga	0	7	1	1
46	Hiburan	Olahraga	0	7	1	1
47	Hiburan	Ekonomi	0	7	1	1
48	Hiburan	Ekonomi	0	7	1	1
49	Olahraga	Olahraga	1	8	0	0
50	Olahraga	Olahraga	1	8	0	0
51	Olahraga	Olahraga	1	8	0	0
52	Olahraga	Olahraga	1	8	0	0
53	Olahraga	Olahraga	1	8	0	0
54	Olahraga	Teknologi	0	7	1	1
55	Olahraga	Teknologi	0	7	1	1

56	Olahraga	Teknologi	0	7	1	1
57	Olahraga	Teknologi	0	7	1	1
58	Olahraga	Teknologi	0	7	1	1
59	Olahraga	Teknologi	0	7	1	1
60	Olahraga	Teknologi	0	7	1	1
61	Pendidikan	Pendidikan	1	8	0	0
62	Pendidikan	Pendidikan	1	8	0	0
63	Pendidikan	Pendidikan	1	8	0	0
64	Pendidikan	Pendidikan	1	8	0	0
65	Pendidikan	Pendidikan	1	8	0	0
66	Pendidikan	Pendidikan	1	8	0	0
67	Pendidikan	Hiburan	0	7	1	1
68	Pendidikan	Hiburan	0	7	1	1
69	Pendidikan	Hiburan	0	7	1	1
70	Pendidikan	Hiburan	0	7	1	1
71	Pendidikan	Hukum	0	7	1	1
72	Pendidikan	Hukum	0	7	1	1
73	Ekonomi	Ekonomi	1	8	0	0
74	Ekonomi	Ekonomi	1	8	0	0
75	Ekonomi	Ekonomi	1	8	0	0
76	Ekonomi	Ekonomi	1	8	0	0
77	Ekonomi	Ekonomi	1	8	0	0
78	Ekonomi	Teknologi	0	7	1	1
79	Ekonomi	Teknologi	0	7	1	1
80	Ekonomi	Teknologi	0	7	1	1
81	Ekonomi	Teknologi	0	7	1	1
82	Ekonomi	Kesehatan	0	7	1	1
83	Ekonomi	Kesehatan	0	7	1	1
84	Ekonomi	Kesehatan	0	7	1	1
85	Hukum	Hukum	1	8	0	0
86	Hukum	Hukum	1	8	0	0
87	Hukum	Hukum	1	8	0	0
88	Hukum	Hukum	1	8	0	0
89	Hukum	Hukum	1	8	0	0
90	Hukum	Politik	0	7	1	1
91	Hukum	Politik	0	7	1	1
92	Hukum	Politik	0	7	1	1
93	Hukum	Pendidikan	0	7	1	1
94	Hukum	Pendidikan	0	7	1	1
95	Hukum	Pendidikan	0	7	1	1

96	Hukum	Pendidikan	0	7	1	1
97	Kesehatan	Kesehatan	1	8	0	0
98	Kesehatan	Kesehatan	1	8	0	0
99	Kesehatan	Kesehatan	1	8	0	0
100	Kesehatan	Kesehatan	1	8	0	0
101	Kesehatan	Kesehatan	1	8	0	0
102	Kesehatan	Kesehatan	1	8	0	0
103	Kesehatan	Olahraga	0	7	1	1
104	Kesehatan	Olahraga	0	7	1	1
105	Kesehatan	Olahraga	0	7	1	1
106	Kesehatan	Olahraga	0	7	1	1
107	Kesehatan	Olahraga	0	7	1	1
108	Kesehatan	Ekonomi	0	7	1	1
Total			51	807	57	57