

Klasifikasi Kepakaran Reviewer Menggunakan Jaringan Syaraf Tiruan (*Artificial Neural Network*)

Akhsinatul Laeliah, Imam Much Ibnu Subroto, Sam Farisa Chaerul Haviana

Teknik Informatika Fakultas Teknologi Industri, Universitas Islam Sultan Agung

Correspondence Author: imam@unissula.ac.id

Abstract

Reviewer jurnal adalah seseorang yang ditugaskan untuk meninjau atau menilai kelayakan suatu artikel untuk dipublikasikan di jurnal tertentu. Seorang reviewer harus memahami isi dari sebuah artikel jurnal yang akan direview sehingga harus ada kesesuaian antara bidang ilmu reviewer dan bidang ilmu yang direview. Kemendikbud sendiri mempunyai sebuah program untuk meningkatkan jumlah publikasi di Indonesia yang dapat bereputasi Internasional. Untuk menjalankan program tersebut dibutuhkan tenaga reviewer yang sesuai dengan bidang ilmu para peneliti dan dosen yang akan publikasi. Dengan banyaknya bidang ilmu yang berbeda-beda maka diperlukan suatu metode untuk mengklasifikasi secara otomatis sehingga pemasaran reviewer dan para peneliti lainnya sesuai. Penelitian ini menggunakan metode *Artificial Neural Network Backpropagation* untuk mengklasifikasikan data reviewer ke 5 bidang utama yaitu *Arts & Humanities, Engineering & Technology, Life Sciences & Medicine, Natural Sciences, dan Social Sciences & Management*. Data reviewer akan melalui proses *preprocessing* teks, kemudian hasil dari *preprocessing* akan dilanjut menggunakan metode *ANN Backpropagation*, dan terakhir untuk mengukur tingkat keakurasian atau kecocokan menggunakan pengukuran akurasi, *precision*, *recall*, *F1-Score* untuk mengukur kinerja klasifikator. Hasil dari pengujian yang dilakukan dengan menggunakan metode *Artificial Neural Network* mendapatkan nilai yang cukup baik yaitu dengan nilai akurasi 0,767 atau setara dengan 76% dengan waktu *testing* dan *training* yang dipunya adalah 47 detik. Selain nilai akurasi juga menghasilkan nilai *recall* 0,566, *precision* 0,570 dan *F1* 0,544.

Keyword: *Artificial Neural Network, Field Area, Reviewer, Data Mining, SINTA*

1. PENDAHULUAN

Review dalam bahasa Indonesia berarti tinjauan atau meninjau, sedangkan reviewer adalah orang yang meninjau suatu hal. Sedangkan reviewer jurnal adalah seseorang yang ditugaskan untuk meninjau dan menilai kelayakan suatu artikel untuk dipublikasikan di jurnal tertentu. Secara ideal reviewer harus memahami isi dari artikel yang direview sehingga harus ada kesesuaian antara bidang ilmu reviewer dan bidang ilmu yang direview. Kemendikbud membuat suatu program untuk meningkatkan jumlah publikasi di Indonesia yang bereputasi Internasional. Untuk menjalankan program tersebut dibutuhkan tenaga reviewer sesuai dengan bidang ilmu para peneliti dan dosen yang akan publikasi. Dengan adanya jumlah reviewer nasional lebih dari 100.000 dengan bidang ilmu yang berbeda-beda, maka diperlukan suatu metode untuk mengklasifikasikan secara otomatis sehingga bidang pemasaran reviewer dan para peneliti lainnya sesuai. Karena adanya permasalahan di KEMENDIKBUD, maka klasifikasi teks menjadi solusi yang tepat karena informasi yang masuk ke dalam SINTA kemudian diklasifikasikan ke dalam kategori tertentu. Beberapa metode yang biasa digunakan dalam klasifikasi teks diantaranya yaitu *Naïve Bayes Classifier (NBC)*, *Support Vector Machine (SVM)*, dan *Neural Network*. Sebelumnya terdapat contoh penelitian tentang klasifikasi teks dengan metode *Artificial Neural Network*. Hasil dari penelitian tersebut mendapat kesimpulan bahwa kategori yang jumlah terbanyak ada pada pendidikan sebanyak 2348 dari total 6916 aduan. 62,2% dari total aduan yang sudah ditindak lanjuti oleh dinas yang berwenang. (1)

Dalam program Kemendikbud ini mengusulkan menggunakan metode ANN (*Artificial Neural Network Backpropagation*) dimana output dari jaringan dibandingkan dengan target yang diharapkan untuk mendapatkan *output error*, kemudian *error* ini dipropagasi kembali untuk memperbaiki bobot jaringan dan meminimalisir *error*. Pada penelitian ini bertujuan untuk pengklasifikasian bidang ilmu setiap artikel yang terindeks di SINTA dan terindeks oleh Scopus yang efektif menjadi 5 bidang ilmu utama yaitu *Arts & Humanities, Engineering & Technology, Life Sciences & Medicine, Natural Sciences, dan Social Sciences & Management*. Pada penelitian sistem klasifikasi kepakaran reviewer menggunakan ANN keberhasilan tergantung pada data-data yang diberikan selama fase penelitian.

Penelitian yang dilakukan oleh Euis Saraswati tentang penerapan Algoritma *Artificial Neural Network* untuk Klasifikasi Opini Publik Terhadap Covid-19 yang mengandung sentimen positif dan negatif menjadikan sebuah pertimbangan dan evaluasi oleh pemerintah dalam menangani virus Covid-19. Pada penelitian ini dengan

menggunakan metode *Artificial Neural Network* dan *Backpropagation* diperoleh nilai akurasi 88,62%, presisi 91,5%, dan *recall* 95,73%. Hasil ini menunjukkan bahwa model ANN cukup baik untuk klasifikasi *text mining*. (2)

2. METODE PENELITIAN

Adapun tahapan dari pengumpulan data untuk menyelesaikan penelitian ini adalah :

1. Data berasal dari SINTA KEMENDIKBUD untuk dapat melakukan sebuah klasifikasi menggunakan metode ANN. Jumlah data yang akan digunakan pada penelitian ini ada 1000 judul artikel dari kategori bidang ilmu.
2. Labelisasi Data
Tahap ini data akan dilabeli sesuai dengan 5 bidang yang sudah ditentukan yaitu *Art & Humanities*, *Engineering & Technology*, *Life Sciences & Medicine*, *Natural Sciences*, dan *Social Sciences & Management*. Setiap judul adekan divalidasi bidang ilmunya oleh pakar yang menguasai klasifikasi bidang ilmu.
3. Sampel data judul artikel dan sampel data *author*
Berikut adalah beberapa contoh data jurul artikel dan data *author* yang peroleh dari data SINTA yang akan di tunjukkan pada tabel 1 dan tabel 2.

Tabel 1 Data Judul

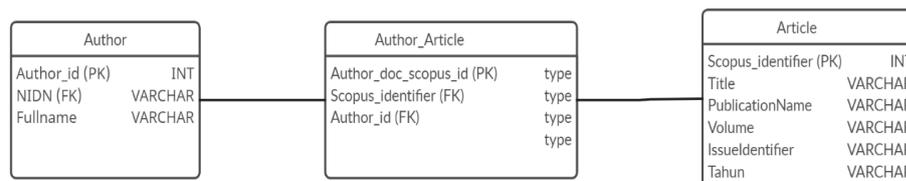
No	Judul	Bidang Ilmu
1.	<i>Analytical Assessments of Foreign Students Writing</i>	<i>Art and Humanities</i>
2.	<i>Asian labor migration ti the Middle East</i>	<i>Art and Humanities</i>
3.	<i>Cancer of the anterior commisure of the larynx</i>	<i>Life Science and Medicine</i>
4.	<i>Occupational hazard in cancer chemotherapy</i>	<i>Life Science and Medicine</i>
5.	<i>Accounting innovation and incentives</i>	<i>Social Science and Management</i>
6.	<i>The best bussiness practices of Sri Lankan firms</i>	<i>Social Science and Management</i>
7.	<i>Sol Gel Materials in Electrochemistry</i>	<i>Natural Science</i>
8.	<i>Photochemistry of some organosulfur pesticides</i>	<i>Natural Science</i>
9.	<i>Computer simulation of turbulence in internal combustion engines</i>	<i>Engineering and Technology</i>
10.	<i>Computer assisted land cover mapping with SPOT in Indonesia</i>	<i>Engineering and Technology</i>

Tabel 2 Data Author

No	Id Author	NIDN	Nama Author
1.	1	0621XXXXXX	ANIS MASHDUROHATUN
2.	28	0603XXXXXX	HERU SULISTYO
3.	31	0627XXXXXX	WIDIYANTO
4.	39	0631XXXXXX	MUHAMMAD QOMARUDDIN
5.	41	0603XXXXXX	ANDRE SUGIYONO
6.	42	0628XXXXXX	ARIEF MASWANTO
7.	44	0613XXXXXX	IMAM MUCH IBNU SUBROTO
8.	45	0605XXXXXX	ANTONIUS
9.	47	0627XXXXXX	PRATIKSO
10.	50	0613XXXXXX	SLAMET IMAM WAHYUDI

2.1 Desain Tabel Database

Database yang digunakan pada pembuatan sistem ini adalah *database* mysql yang berisi beberapa tabel antara lain : tabel dataset, tabel reviewer, tabel judul, tabel artikel. *Database* tersebut kemudian akan diolah dengan beberapa *query* untuk menampilkan data sesua kebutuhan sistem. Dalam proses pengolahan data membutuhkan relasi data antar tabel. Untuk relasi antar tabel yang dibutuhkan dalam sistem akan tunjukkan pada gambar 2.1.



Gambar 1 Desain Tabel Database

Pada gambar 1 merupakan relasi dari *database* yang memiliki nama tabel dataset, tabel reviewer, tabel judul, tabel artikel. Untuk melakukan relasi antar tabel menggunakan *foreign key* (kunci tamu).

2.2 Preprocessing

Tahap ini adalah melakukan identifikasi atau melakukan pengkodean dalam metode ANN untuk untuk merubah data dari data mentah menjadi data dalam format yang berguna dan efisien. Tahapan *Preprocessing* yang digunakan yaitu : data *cleaning*, *case folding*, *stopwords removal*, *stemming*, dan *tokenizing*. (3) (4)

- Data Cleaning*, yaitu membersihkan data teks dari kata-kata yang tidak diperlukan untuk mengurangi noise.
 - Case Folding*, yaitu proses mengubah kata menjadi bukan kapital dan menghilangkan tanda baca maupun angka.
 - Stopwords Removal*, yaitu proses menghilangkan kata-kata yang tidak memiliki arti yang relevan.
 - Stemming*, yaitu proses menghilangkan imbuhan pada kata untuk mendapat kata kunci dan kata dasar.
 - Tokenizing*, yaitu proses pemotongan teks menjadi potongan yang lebih kecil, atau disebut token.
- (1) (5)

2.3 Penerapan Metode ANN

Tahap ini adalah melakukan penerapan metode pada penelitian yaitu menggunakan metode *Artificial Neural Network* terhadap dataset yang dihasilkan dari *Preprocessing*. Model yang digunakan ANN adalah *Backpropagation*.

Pada ANN terdapat turunan fungsi aktivasi yang memberikan *output* berdasarkan dari sinyal *input*. Tujuan utama dari fungsi aktivasi adalah menghasilkan nilai *output* berdasarkan input yang ada dan dilakukan dalam lapisan tersembunyi. Terdapat beberapa turunan fungsi aktivasi seperti berikut : (6)

Turunan fungsi aktivasi

a. Sigmoid

$$y = f(x) = \frac{1}{1 + e^{-\alpha x}} \quad (1)$$

```
// ALGORITMA KLASIFIKASI
], new MultilayerPerceptron([
  new Dense(60),
  new Activation(new Sigmoid()),
  new Dense(50),
  new Activation(new Sigmoid()),
  new Dense(25),
]),
```

Gambar 2 Metode Artificial Neural Network di Rubix MI

2.4 Evaluasi

Tahap ini adalah mengevaluasi data yang sudah diproses sebelumnya menggunakan metode *Backpropagation* sehingga menghasilkan data yang sesuai kategori yang ditentukan. Evaluasi pada penelitian ini menggunakan pengukuran akurasi, *precision*, *recall* untuk mengukur kinerja klasifikator.

Akurasi merupakan rasio untuk mengukur kebenaran dalam keseluruhan data.

$$\text{Akurasi} = (TP + TN) / (TP + FP + FN + TN) \quad (2)$$

Precision merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif.

$$\text{Precision} = (TP) / (TP + FP) \quad (3)$$

Recall merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif.

$$\text{Recall} = (TP) / (TP + FN) \quad (4)$$

F1 Score merupakan perbandingan rata-rata presisi dan *recall* yang dibobotkan.

$$F1 \text{ Score} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}) \quad (5)$$

Keterangan :

TP = *True Positive*

TN = *True Negative*

FP = *False Positive*

FN = *False Negative*

3. HASIL DAN ANALISA (10 PT)

Pada bagian ini, dijelaskan hasil penelitian dan pada saat yang sama diberikan pembahasan yang komprehensif. Hasil dapat disajikan dalam angka, grafik, tabel dan lain-lain yang membuat pembaca memahami dengan mudah [2], [5]. pembahasan dapat dibuat dalam beberapa sub-bab.

3.1 Pengujian Sistem

Pada pengujian sistem ini menggunakan metode pengujian *Black Box Testing*, dimana *Black Box Testing* merupakan sebuah pengujian pada perangkat lunak yang bertujuan untuk melakukan pengetesan fungsionalitas aplikasi tanpa perlu melihat ke dalam struktur atau cara kerja internalnya. Pada pengujian sistem ini dilakukan untuk mengetes setiap fungsi yang ada pada sistem.

Tabel 3 Pengujian Sistem

Skenario Pengujian	Kasus Pengujian	Hasil Pengujian	Keterangan
<i>Upload</i> Data Judul	Melakukan <i>Upload</i> data judul dari data 5 bidang ilmu	Sesuai	Sukses
<i>Upload</i> Data <i>Author</i>	Melakukan <i>Upload</i> data <i>author</i> yang sudah dipilih	Sesuai	Sukses
<i>Author</i> memiliki beberapa Judul	Melakukan <i>testing</i> judul yang dimiliki <i>author</i>	Sesuai	Sukses
Judul di prediksi ke 5 bidang ilmu	Melakukan <i>testing</i> judul sesuai dengan 5 bidang ilmu yang sudah ditentukan.	Sesuai	Sukses

Dari hasil pengujian yang dilakukan dapat dilihat bahwa sistem prediksi judul ke dalam 5 bidang ilmu berbasis web sudah sesuai dengan setiap fungsi yang dimiliki.

3.2 Analisis Hasil Akurasi

Selanjutnya adalah Tahap Analisis Hasil Akurasi yang bertujuan untuk mengetahui seberapa tingkat akurasi yang telah dibuat ini, tahap ini akan menggunakan data *testing* dengan pembagian dari data *testing* dan data sampel.

Tabel 4 Data Training dan Data Testing

Nama Data	Jumlah (%)	Jumlah Angka
Data Training	70 %	350
Data Testing	30 %	150

Total data : 500

Tabel 5 Data dibagi ke 5 bidang ilmu

Class	<i>Art & Humanities</i>	<i>Engineering & Technology</i>	<i>Life Science & Medicine</i>	<i>Natural Science</i>	<i>Social Science & Management</i>
Data Training	70	70	70	70	70
Data Testing	30	30	30	30	30

Data diatas nantinya akan digunakan untuk perhitungan akurasi yang menghasilkan sebuah confusional matrix untuk mencari nilai *accuracy*, *precision*, *recall* dan juga F1 dengan berbagai varian jumlah *dense* (node), yaitu (60, 50, 25), (70, 50, 25), (80, 50, 25) (90, 50, 25), (100, 50, 25), (110, 50, 25), (120, 50, 25), (125, 50, 25), (130, 50, 25), (150, 50, 25) yang nantinya akan dicari nilai *dense* terbaik dalam melakukan prediksi judul kedalam 5 bidang ilmu.

Tabel 6 Dense (60, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	85 TP	65 FP
<i>Actually Negative (I)</i>	65 FN	340 TN

Tabel 7 Dense (70, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	69 TP	81 FP
<i>Actually Negative (I)</i>	81 FN	276 TN

Tabel 8 Dense (80, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	83 TP	67 FP
<i>Actually Negative (I)</i>	67 FN	332 TN

Tabel 9 Dense (90, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	73 TP	77 FP
<i>Actually Negative (I)</i>	77 FN	292 TN

Tabel 10 Dense (100, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	73 TP	77 FP
<i>Actually Negative (I)</i>	77 FN	292 TN

Tabel 11 Dense (110, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	82 TP	68 FP
<i>Actually Negative (I)</i>	68 FN	328 TN

Tabel 12 Dense (120, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	84 TP	66 FP
<i>Actually Negative (I)</i>	66 FN	336 TN

Tabel 13 Dense (125, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	77 TP	73 FP
<i>Actually Negative (I)</i>	73 FN	308 TN

Tabel 14 Dense (130, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	83 TP	67 FP
<i>Actually Negative (I)</i>	67 FN	332 TN

Tabel 15 Dense (150, 50, 25)

	<i>Actually Positive (I)</i>	<i>Actualt Negatif (I)</i>
<i>Predicted Positive (I)</i>	68 TP	82 FP
<i>Actually Negative (I)</i>	82 FN	272 TN

Dengan data confusion matrix yang sudah diperoleh, maka dapat diketahui *accuracy*, *precision*, *recall*, dan F1 dari setiap *dense*. Untuk hasilnya akan diperlihatkan pada tabel 16.

Tabel 16 Hasil Akurasi

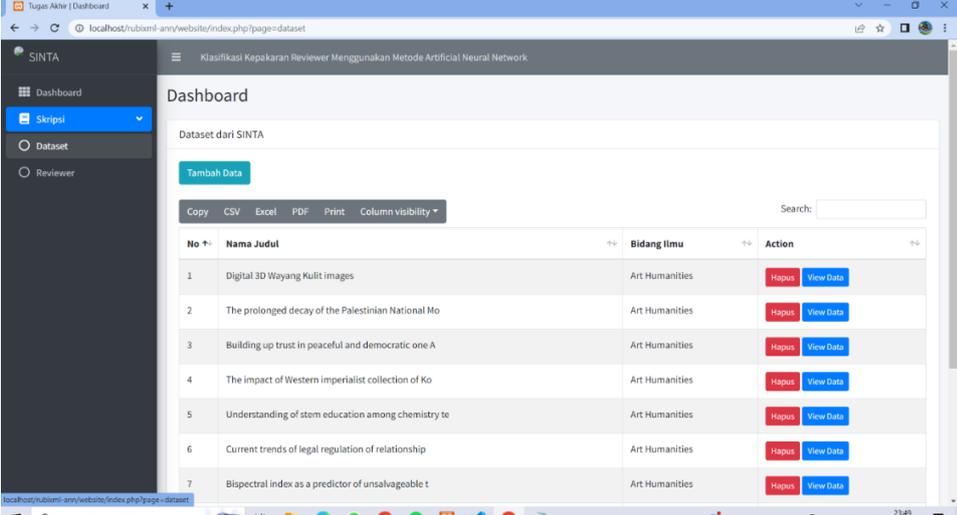
<i>Dense</i>	<i>Accuracy</i>	<i>Recall</i>	<i>Precision</i>	F1	<i>Time</i>
60, 50, 25	0,767	0,566	0,570	0,544	47 detik
70, 50, 25	0,685	0,459	0,503	0,452	1 menit 5 detik
80, 50, 25	0,759	0,553	0,608	0,535	52 detik
90, 50, 25	0,712	0,486	0,607	0,600	1 menit 6 detik
100, 50, 25	0,706	0,486	0,536	0,474	1 menit 27 detik
110, 50, 25	0,753	0,546	0,562	0,529	1 menit 5 detik
120, 50, 25	0,762	0,56	0,550	0,538	1 menit 10 detik
125, 50, 25	0,734	0,513	0,579	0,515	1 menit 51 detik
130, 50, 25	0,757	0,553	0,583	0,536	1 menit 35 detik
150, 50, 25	0,688	0,453	0,564	0,442	1 menit 46 detik

Berdasarkan hasil pengujian yang sudah dilakukan, dat ditarik kesimpulan bahwa dalam pengujian akurasi berdasarkan *dense* yang sudah ditentukan sebelumnya memiliki nilai akurasi yang baik, terutama pada *dense* yang bernilai (60, 50, 25) yang menghasilkan nilai akurasi paling tinggi dari pada nilai *dense* lain dengan nilai akurasi

0,767 atau setara dengan 76% dengan waktu *testing* dan *training* yang dipunya adalah 47 detik. Untuk nilai *dense* (60, 50, 25) juga memiliki nilai *precision*, *recall* dan F1 yang cukup tinggi dari pada *dense* lain yaitu *recall* 0,566, *precision* 0,570 dan F1 0,544. Untuk nilai *dense* lain nilai akurasi lebih rendah diantaranya 0,685 sampai dengan 0,762.

3.3 Implementasi User Interface (UI)

Implementasi merupakan tahapan setelah rancangan sistem selesai dibuat, pada sistem ini sudah dibuat rancangan tampilan *User Interface* web. Untuk hasil dari implementasi *User Interface* ditunjukkan pada gambar 3.1 sampai 3.3 :



Dashboard

Dataset dari SINTA

Tambah Data

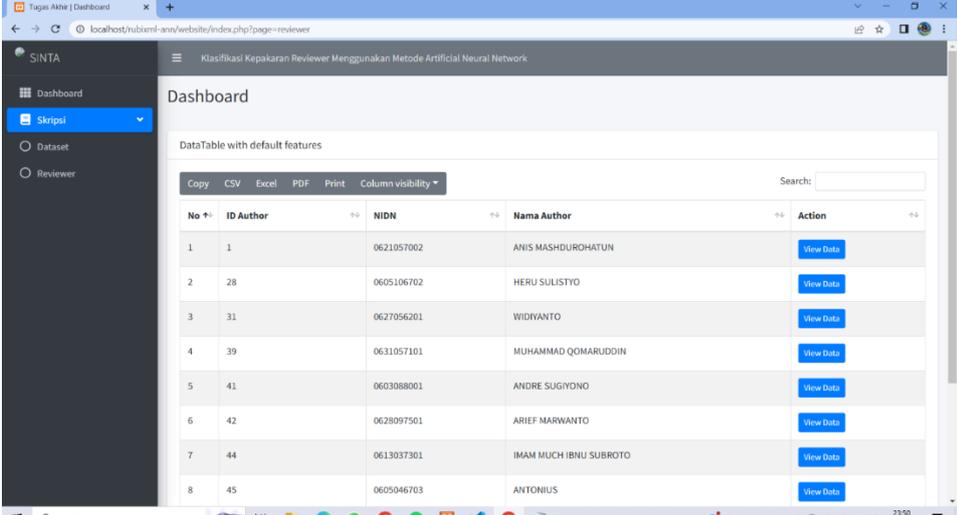
Copy CSV Excel PDF Print Column visibility

Search:

No	Nama Judul	Bidang Ilmu	Action
1	Digital 3D Wayang Kulit images	Art Humanities	Hapus View Data
2	The prolonged decay of the Palestinian National Mo	Art Humanities	Hapus View Data
3	Building up trust in peaceful and democratic one A	Art Humanities	Hapus View Data
4	The impact of Western imperialist collection of Ko	Art Humanities	Hapus View Data
5	Understanding of stem education among chemistry te	Art Humanities	Hapus View Data
6	Current trends of legal regulation of relationship	Art Humanities	Hapus View Data
7	Bispectral index as a predictor of unsalvageable t	Art Humanities	Hapus View Data

Gambar 3 UI dari Dataset

Pada gambar 3 merupakan halaman dari Dataset yang berisikan data judul dengan bidang ilmu yang dimiliki. Pada halaman ini juga terdapat tombol tambah data yang berfungsi untuk apabila ingin menambahkan data judul serta tombol download file sesuai dengan tipe yang diinginkan. Selain itu, terdapat juga tombol 'hapus' yang berfungsi untuk menghapus beberapa data yang diinginkan.



Dashboard

DataTable with default features

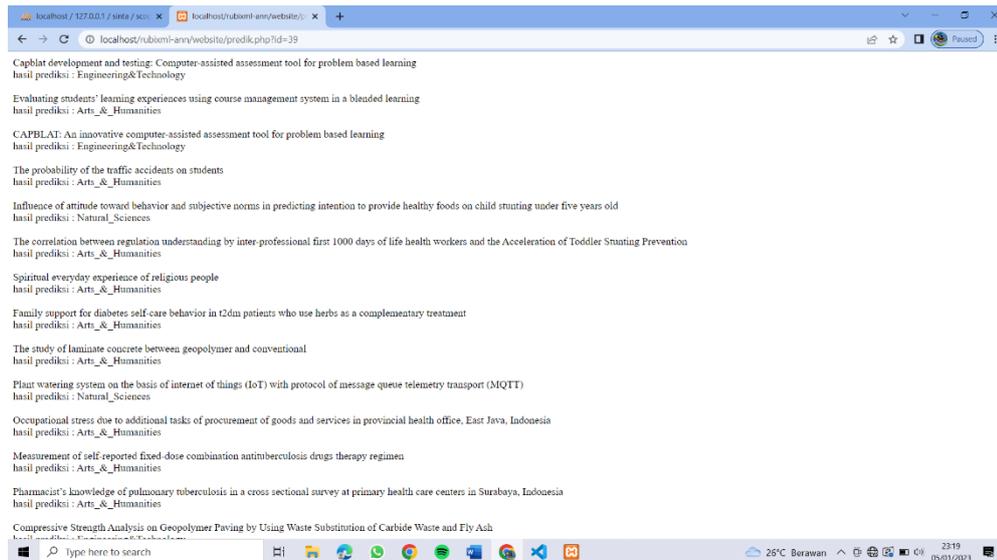
Copy CSV Excel PDF Print Column visibility

Search:

No	ID Author	NIDN	Nama Author	Action
1	1	0621057002	ANIS MASHDUROHATUN	View Data
2	28	0605106702	HERU SULISTYO	View Data
3	31	0627056201	WIDVANTO	View Data
4	39	0631057101	MUHAMMAD QOMARUDDIN	View Data
5	41	0603088001	ANDRE SUGIYONO	View Data
6	42	0628097501	ARIEF MARWANTO	View Data
7	44	0613037301	IMAM MUCH IBNU SUBROTO	View Data
8	45	0605046703	ANTONIUS	View Data

Gambar 4 UI dari Data Reviewer Author

Pada gambar 4 merupakan tampilan dari halaman Data Reviewer yang berisi data *author* yang memiliki atribut *ID Author*, *NIDN*, *Nama Author*. Pada halaman ini juga terdapat tombol untuk mendownload data reviewer sesuai dengan tipe text yang diinginkan. Selain itu, pada halaman ini juga terdapat tombol aksi yang bernama 'view *author*' yang berfungsi untuk melihat jumlah data judul yang dimiliki oleh salah satu dari *author*.



Gambar 5 UI dari Data Reviewer Prediksi

Pada gambar 5 merupakan tampilan halaman dari Data Reviewer Prediksi yang menampilkan hasil dari prediksi judul ke dalam 5 bidang ilmu dan menampilkan persentase melalui chart dari setiap bidang ilmu.

4. KESIMPULAN

Penulis dapat menyimpulkan bahwa hasil dari pengujian yang dilakukan dengan menggunakan metode *Artificial Neural Network* memperoleh hasil yang cukup baik yaitu dengan nilai akurasi 0,767 atau setara dengan 76% dengan waktu *testing* dan *training* yang dipunya adalah 47 detik. Selain nilai akurasi juga menghasilkan nilai *recall* 0,566, *precision* 0,570 dan *F1* 0,544.

ACKNOWLEDGEMENTS

Tugas akhir ini merupakan bagian dari penelitian dosen Teknik Informatika Universitas Islam Sultan Agung berjudul Pengembangan Sistem Klasifikasi Otomatis Publikasi Ilmiah pada database SINTA Science and Technology Index yang dibiayai dari Kementerian Pendidikan dan Kebudayaan, Pendidikan Tinggi, Riset dan Teknologi Republik Indonesia tahun 2022.

DAFTAR PUSTAKA

- [1]. Ananto MI, Winahju WS, Fithriasari K. Klasifikasi Kategori Pengaduan Masyarakat Melalui Kanal LAPOR! Menggunakan Artificial Neural Network. *Inferensi*. 2019;2(2):71.
- [2]. Euis Saraswati, Yuyun Umaidah, Apriade Voutama. Penerapan Algoritma Artificial Neural Network untuk Klasifikasi Opini Publik Terhadap Covid-19. *Gener J*. 2021;5(2):109–18.
- [3]. Harjanta ATJ. Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining. *Inform UPGRIS*. 2015;1:1–9.
- [4]. Ma'rifah H, Wibawa AP, Akbar MI. Klasifikasi Artikel Ilmiah Dengan Berbagai Skenario Preprocessing. *Sains, Apl Komputasi dan Teknol Inf*. 2020;2(2):70.
- [5]. Hermawan L, Bellanier Ismiati M. Pembelajaran Text Preprocessing berbasis Simulator Untuk Mata Kuliah Information Retrieval. *J Transform*. 2020;17(2):188.
- [6]. Widodo AP, Sarwoko EA, Firdaus Z. Akurasi Model Prediksi Metode Backpropagation Menggunakan Kombinasi Hidden Neuron Dengan Alpha. *Matematika*. 2017;20(2):79–84.