

Deteksi Awal Kelulusan Mahasiswa Dengan Algoritma Jaringan Syaraf Tiruan Dan Information Gain

Hafiyy Viazola Aziz¹, Dr. Kusuma Ayu Laksitowening², Ibnu Asror³

^{1,2,3} Universitas Telkom, Bandung

¹viazola@students.telkomuniversity.ac.id, ²ayu@telkomuniversity.ac.id, ³iasror@telkomuniversity.ac.id

Abstrak

Persentase kelulusan tepat waktu mahasiswa menjadi salah satu poin penilaian oleh Badan Akreditasi Nasional Perguruan Tinggi yang akan berpengaruh pada nilai akreditasi perguruan tinggi. Penelitian ini mengidentifikasi pentingnya mendeteksi kelulusan mahasiswa seawal mungkin sehingga bisa dilakukan tindakan preventif untuk mahasiswa yang terdeteksi kelulusannya tidak tepat waktu, baik itu dari pihak perguruan tinggi maupun individu. Dengan adanya penelitian ini, diharapkan persentase kelulusan tepat waktu mahasiswa Fakultas Informatika Universitas Telkom akan meningkat. Deteksi dini ini dapat dilakukan dengan teknik *data mining* yaitu klasifikasi, klasifikasi dapat dilakukan dengan berbagai teknik maupun algoritma, di antaranya dengan algoritma Jaringan Syaraf Tiruan dan *Information Gain*. Algoritma jaringan syaraf tiruan akan menghasilkan nilai akurasi berdasarkan atribut yang diperoleh dari nilai *Information Gain*, sedangkan *Information Gain* akan mencari nilai entropy dari setiap atribut sehingga diketahui atribut yang berpengaruh untuk diproses di algoritma jaringan syaraf tiruan yang akan menghasilkan akurasi yang optimal. Hasil dari penelitian ini berupa atribut yang paling berpengaruh berdasarkan *Information Gain* dan akurasi prediksi dengan algoritma jaringan syaraf tiruan. Atribut yang dipakai berjumlah 60% dari total atribut yang menghasilkan akurasi 74.5% untuk tingkat 1, 78.6% untuk tingkat 2, 81.9% untuk tingkat 3, dan 83.6% untuk tingkat 4.

Kata kunci : *educational data mining*, Klasifikasi, *Information gain*, Jaringan syaraf tiruan

Abstract

The percentage of students graduating on time is one of the points of assessment by the National Accreditation Board for Higher Education which will have an effect on the college accreditation score. This study identifies the importance of detecting student graduation as early as possible so that preventive action can be taken for students who are detected that their graduation is not on time, either from the university or individuals. With this research, it is expected that the percentage of on-time graduation of students from the Faculty of Informatics, Telkom University will increase. This early detection can be done with data mining techniques, namely classification, classification can be done with various techniques and algorithms, including the Artificial Neural Network algorithm and *Information Gain*. The artificial neural network algorithm will produce an accuracy value based on the attributes obtained from the *Information Gain* value. *Information Gain* will look for the entropy value of each attribute so that the influential attributes are known to be processed in the neural network algorithm which will produce optimal accuracy. The results of this study are the most influential attributes based on *Information Gain* to be processed in the artificial neural network algorithm. The attributes used are 60% of the total attributes which result in an accuracy of 74.5% for level 1, 78.6% for level 2, 81.9% for level 3, and 83.6% for level 4.

Keyword : Educational Data mining, Classification, Information gain, Artificial Neural Network

1. Pendahuluan

Persentase kelulusan mahasiswa menjadi faktor penting yang perlu diperhatikan oleh setiap perguruan tinggi yang ada di Indonesia. Elemen penilaian pada lulusan perguruan tinggi salah satunya adalah perguruan tinggi memiliki angka efisien edukasi yang ideal (BAN PT, 2007). Banyak faktor yang mempengaruhi persentase kelulusan mahasiswa, diantaranya belum tercukupinya SKS untuk lulus dan nilai mata kuliah yang belum memenuhi syarat. Waktu kelulusan mahasiswa tidak bisa diprediksi secara instan, maka dibutuhkan teknik yang secara spesifik dan efektif bisa memprediksi kelulusan mahasiswa dengan parameter-/parameter yang tepat. Teknik yang selama ini umum digunakan untuk memprediksi adalah *Data Mining* yang nantinya akan menggunakan metode klasifikasi.

Data mining adalah analisis terhadap data (biasanya data yang berukuran besar) untuk menemukan keterkaitan yang jelas serta menyimpulkannya yang nantinya akan berguna bagi pemilik data tersebut[1]. *Data mining* mempunyai banyak metode maupun teknik sehingga sangat fleksibel digunakan untuk banyak kasus yang berkaitan dengan prediksi[2]. Model algoritma data mining yang umum digunakan yaitu jaringan syaraf tiruan, naïve bayes, Decision Tree, K-Nearest Neighbor dan Logistic Regression. Dalam penelitian ini akan dilakukan pengujian memakai jaringan syaraf tiruan dan pemilihan atribut memakai *information gain*.

Algoritma jaringan syaraf tiruan,yaitu algoritma yang mengadopsi mekanisme berpikir sebuah sistem berdasarkan otak manusia[3]. Algoritma jaringan syaraf tiruan kurang optimal untuk proses numerik maka dari itu dibutuhkan information gain sebagai feature selection untuk memaksimalkan hasil pengujian.

Information Gain adalah proses perubahan *entropy* yang terjadi setelah mempartisi data berdasarkan atribut untuk mendapatkan atribut terpilih yang akan digunakan dalam proses prediksi. Tujuan dari information gain adalah menampilkan seluruh atribut beserta dengan nilai information gain nya untuk mengetahui atribut terbaik untuk proses prediksi.Information gain akan dikombinasikan dengan algoritma jaringan syaraf tiruan untuk mendapatkan hasil prediksi.

2. Tinjauan Pustaka

2.1 Studi Terkait

Pada penelitian [4] yang dilakukan oleh Utomo Budiyanto dengan judul “Prediksi Kelulusan Tepat Waktu Menggunakan Jaringan Syaraf Tiruan”.Dalam penelitian ini,peneliti menggunakan data wisudawan semester genap tahun 2017 dan semester gasal 2018.Langkah yang dilakukan oleh peneliti diantaranya adalah proses pengumpulan Data, pemisahan data untuk training dan testing, menentukan struktur jaringan, menentukan algoritma Pembelajaran serta uji keluaran.

Penelitian selanjutnya [5] yang dilakukan oleh Eko Prasetyo dengan judul “Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree Dan Artificial Neural Network”,tahap pertama dalam penelitian tersebut adalah pengumpulan data akademik mahasiswa dan *preprocessing* data untuk menghasilkan atribut yang *fixed*. Nantinya data akan diolah dengan 2 metode yaitu *Decision Tree* dan *Artificial Neural Network*. Dalam penelitian ini akurasi yang didapat oleh *artificial neural network* lebih besar dari *decision tree* yaitu 79,74% berbanding 74,51%.Kelebihan penelitian ini terletak pada perbandingan kedua algoritma yang digunakan dan penulis menjelaskan secara detail bagaimana proses dari awal hingga hasil prediksi keluar.Ada perbedaan pada penelitian [5] dengan penelitian ini,penelitian ini akan menggunakan pemilihan atribut terlebih dahulu sebelum algoritma *Artificial Neural Network* berjalan sehingga berpengaruh kepada hasil akurasi.

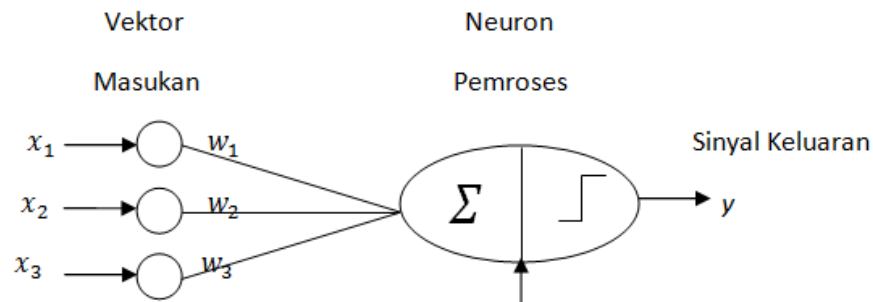
Penelitian selanjutnya berkaitan dengan klasifikasi memakai artificial neural network yang berjudul “*Fully automatic model-based segmentation and classification approach for MRI brain tumor using artificial neural networks*”. Pada penelitian mempresentasikan pendekatan segmentasi dan klasifikasi berbasis model robot otomatis untuk tumor otak MRI menggunakan jaringan saraf tiruan.Hasilnya telah menunjukkan metode strategi baru yang mengotomatiskan segmentasi tumor otak dan mengidentifikasinya dari gambar MRI. Strategi tersebut menggunakan proses segmentasi dan pendekatan *Trained Ann* untuk secara tepat mengenali lokasi ROI.Gambar yang dihasilkan disusun menggunakan JST untuk mendapatkan tumor tersegmentasi. Hasil akurasi yang didapat pada penelitian kali ini sebesar 92.14%[6].

2.2 Educational Data Mining

Educational Data Mining (EDM) adalah pengaplikasian teknik data mining dalam dunia pendidikan. Proses EDM menjadikan guru dan peneliti pendidikan mendapat informasi dengan mengolah data mentah menjadi informasi yang dapat dipahami[7]. Selain itu keuntungan besar dari penggunaan metode EDM adalah dapat dimasukan sebagai bagian dari pengembangan system e-learning. Dengan EDM dapat ditemukan pengetahuan baru dari penggalian data siswa, sehingga berpotensi meningkatkan beberapa aspek kualitas pendidikan [8].Fungsi data mining yang akan digunakan pada penelitian kali ini adalah fungsi klasifikasi, algoritma yang digunakan dalam proses klasifikasi adalah algoritma jaringan syaraf tiruan.

2.3 Jaringan Syaraf Tiruan

Algoritma yang akan digunakan adalah algoritma jaringan syaraf tiruan,yaitu merupakan model penalaran yang didasarkan pada otak manusia.Jaringan Syaraf Tiruan terdiri dari sejumlah prosesor sangat sederhana dan saling berhubungan yang disebut neuron. Neuron yang terhubung dengan pembobotan (weight) melewati sinyal dari neuron satu ke neuron yang lain[9]. Artificial Neural Network meniru cara kerja otak manusia sehingga bisa menyelesaikan permasalahan dengan cara *learning*. Artificial Neural Network memiliki kelebihan diantaranya mampu belajar dari data masukan yang diberikan hingga bisa memetakan hubungan antara masukan dan keluarannya [10]. Klasifikasi menggunakan Artificial Neural Network cukup banyak digunakan oleh para peneliti karena mampu melakukan generalisasi dan ekstraksi dari suatu pola data tertentu serta ANN dapat menciptakan suatu pola pengetahuan melalui pengaturan diri atau kemampuan belajar (self organizing) [11]. Lalu feature selection atau pemilihan atribut yang digunakan adalah information gain,information gain adalah Teknik pemilihan atribut berdasarkan nilai entropy dari setiap atribut yang ada pada dataset.Information gain digunakan untuk mengetahui atribut mana yang paling berpengaruh terhadap kelulusan mahasiswa.

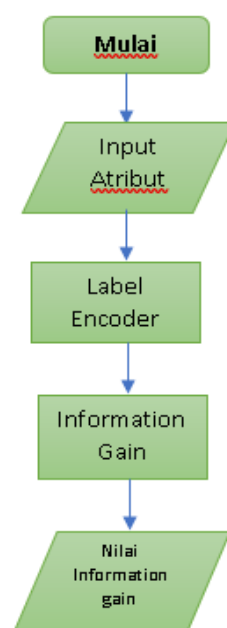


Gambar 1. Model Jaringan Syaraf Tiruan

Pada gambar diatas, dapat dilihat bahwa vektor masukan terdiri dari sejumlah fitur yang diberikan sebagai nilai masukan pada jaringan syaraf tiruan, vektor masukan tersebut memiliki 3 nilai (x_1, x_2, x_3) sebagai fitur dalam vektor yang akan diproses dalam jaringan syaraf tiruan, masing-masing nilai masukan melewati sebuah penghubung berlabelkan w , kemudian semua nilai akan digabungkan. Nilai gabungan tersebut kemudian diproses oleh fungsi aktivasi untuk menghasilkan sinyal y sebagai keluaran. Fungsi aktivasi menggunakan sebuah nilai ambang batas untuk membatasi nilai keluaran agar selalu dalam batas nilai yang ditetapkan. Perancangan algoritma jaringan syaraf tiruan, umumnya dibagi menjadi dua proses utama yaitu pelatihan dan pengujian. Sebelum kedua proses itu dilakukan, perlu dipersiapkan pembagian data untuk data latih dan data uji terlebih dahulu. Persentase pembagian data latih dengan data uji yang digunakan yaitu 70:30. Proses pelatihan dilakukan menggunakan sekumpulan data latih yang memuat feature yang digunakan untuk membedakan antara objek satu dengan objek lainnya. Pemodelan yang digunakan dalam penelitian ini adalah multi-layer perceptron, multi-layer perceptron adalah salah satu pemodelan dalam teknologi jaringan saraf tiruan dengan karakteristik memiliki nilai bobot yang lebih baik dari pada pemodelan yang lain, sehingga menghasilkan klasifikasi yang lebih akurat

2.4 Information Gain

Information gain merupakan salah satu metode pemilihan atribut yang akan menghasilkan atribut terbaik berdasarkan nilai entropy. Entropy merupakan distribusi probabilitas dalam teori informasi dan diadopsi kedalam algoritma C4.5 untuk mengukur tingkat homogenitas distribusi kelas dari sebuah himpunan (dataset). Berikut merupakan tahapan dan formula information gain dalam penelitian ini



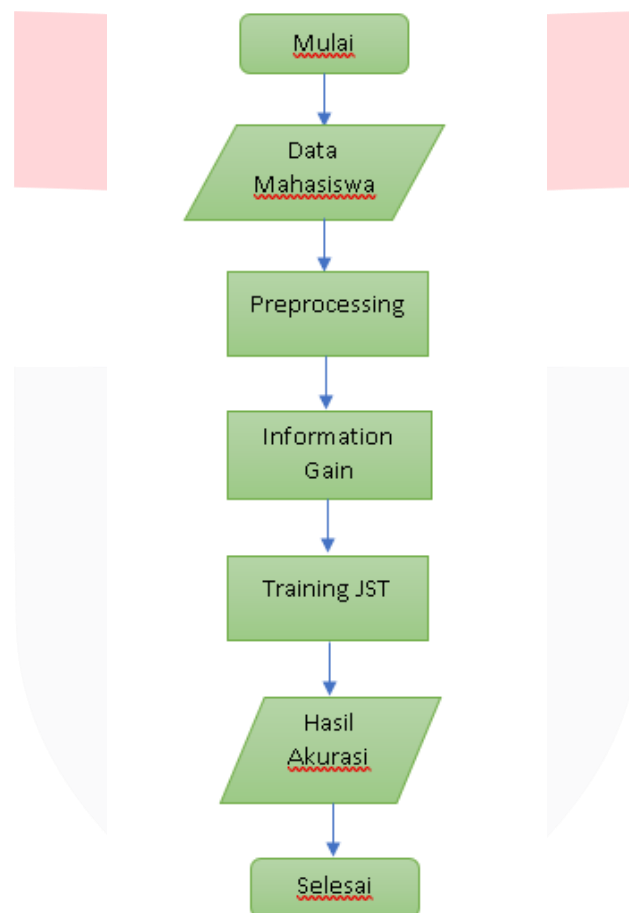
Gambar 2. Flowchart Information Gain

Pada gambar diatas, proses information gain diawali dengan input atribut, lalu *label encoder* untuk memudahkan data dalam proses penghitungan nilai information gain, penghitungan ini memakai library *mutual information*..Setelah semua fitur diproses maka akan keluar hasil berupa nilai information gain untuk semua fitur, hasil itu selanjutnya diurutkan dari nilai information gain terbesar hingga terkecil.

3. Metodologi

3.1 Skema Sistem

Pada penelitian ini, sistem yang dirancang menggunakan algoritma jaringan syaraf tiruan dan information gain untuk deteksi awal kelulusan mahasiswa. Dataset yang digunakan dalam penelitian ini adalah data mahasiswa dan data nilai mahasiswa S1 Teknik Informatika Universitas Telkom tahun 2008-2011 karena parameter dan atribut yang digunakan masih relevan. Berikut merupakan flowchart sistem yang dibangun :



Gambar 3. Flowchart Sistem

Pada flowchart diatas, proses pembangunan sistem akan dimulai dengan input dataset mahasiswa, lalu dataset akan memasuki tahapan preprocessing .Setelah itu, algoritma information gain akan memproses data hasil preprocessing untuk mendapatkan atribut terbaik yang nantinya akan digunakan di tahap *training* JST. Lalu, setelah atribut terbaik diproses oleh algoritma jaringan syaraf tiruan maka keluarlah hasil akurasi persentase kelulusan mahasiswa.

3.2 Preprocessing Data

Preprocessing data akan menggunakan Microsoft Excel dan Pentaho untuk mengeliminasi atribut yang tidak diperlukan dalam pengujian. Atribut yang akan dipakai dalam data mahasiswa adalah NIM, Prodi, Angkatan, Status, Tahun Lulus, dan Lama Kuliah. Sedangkan dalam data nilai hanya memakai atribut NIM, Kode, Mata Kuliah, SKS, Semester, Nilai, Bobot, dan Tingkat. Lalu data akan digabungkan serta membaginya menjadi bertingkat dengan kondisi data sudah difilter sehingga hanya menggunakan data Prodi S1 Teknik

Informatika. Setelah data di filter dan digabung, selanjutnya data duplikat akan dihapus berdasarkan NIM dan Mata Kuliah.

Tabel 1. Data Mahasiswa dan Data Nilai

NIM	PRODI	ANGKATAN	STATUS	Waktu Kuliah	KODE	MATAKULIAH	SKS	SEMESTER	NILAI	BOBOT	TINGKAT
s007e4236c1344b0	S1 Teknik Informatika	2011	GRADUATED	4	F11901	PRAKTIKUM FISIKA	1	2	A	4	1
s007e4236c1344b0	S1 Teknik Informatika	2011	GRADUATED	4	MA1124	KALKULUS II	4	2	B	3	1
s007e4236c1344b0	S1 Teknik Informatika	2011	GRADUATED	4	MA1223	ALJABAR LINEAR	3	2	A	4	1
s01c7afff21068d6	S1 Teknik Informatika	2011	GRADUATED	5	CS1113	PEMROGRAMAN KOMPUTER	3	1	A	4	1
s01c7afff21068d6	S1 Teknik Informatika	2011	GRADUATED	5	CS1911	PRAKTIKUM PEMROGRAMAN KOMPUTER	1	1	A	4	1

Setelah didapatkan data mahasiswa dan data nilai, tahap selanjutnya merubah indeks nilai A menjadi 4, nilai B menjadi 3, nilai C menjadi 2, nilai D menjadi 1, lalu 0 untuk nilai E, missing value, dan nilai T. Jika atribut Mata Kuliah memiliki nilai 0-3 maka akan dilabeli DO, untuk nilai 4 akan dilabeli tepat waktu (TP) dan untuk nilai diatas 4 akan dilabeli telat waktu (TW).

Tabel 2. Data Akhir Mahasiswa 2009 tingkat 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	NIM	AGAMA	A	ALJABAR	I	BAHASA	II	BAHASA	II	FISIKA	INTERNET	KALKULUS	KALKULUS	PEMROGA	PENGANT	PENGETAH	PRAKTIKU	PRAKTIKU
2	s01eb69f5e7a5650		3	2	2	3	3	3	3	2	3	3	2	4	4	4	2.76	TP
3	s01ff3441fee5bd9		4	3	2	3	3	3	4	3	3	3	2	4	4	4	3.11	TP
4	s0254d181a498504		3	0	2	1	1	2	0	0	0	2	2	2	2	3	1.26	TW
5	s03406fad2b25488		3	3	3	2	3	3	4	3	4	3	3	2	3	4	3.13	TP
6	s0393208263e0e60		3	4	4	3	2	3	4	3	4	2	3	2	4	4	3.24	TP
7	s03ffb4af0baad33		3	3	4	3	3	2	4	3	4	3	4	3	3	4	3.27	TW
8	s0495b6264c8a521		3	4	3	3	2	3	3	1	4	3	3	2	3	4	2.89	TW
9	s04f5278ce8ecd0a		3	4	4	3	3	3	4	4	4	4	4	3	3	4	3.54	TW
10	s05622fd10d4c8f5		4	3	2	2	2	3	2	3	3	3	2	2	3	4	2.71	TW
11	s05e6d3c282533af		3	1	2	3	2	2	2	3	1	2	2	2	3	3	2.13	TW
12	s06403f057905f94		3	3	4	4	3	4	3	4	3	4	3	3	3	4	3.24	TW

4. Hasil

Berikut adalah skenario pengujian yang dilakukan sesuai dengan rumusan masalah pada penelitian ini :

1. Pemilihan atribut dari dataset mahasiswa S1 Teknik Informatika Telkom University tahun 2008-2011 sehingga diketahui atribut yang berpengaruh terhadap kelulusan. Pemilihan atribut menggunakan metode Information Gain lalu hasilnya diurutkan berdasarkan atribut dengan nilai Information Gain tertinggi
2. Pengujian prediksi kelulusan mahasiswa menggunakan algoritma jaringan syaraf tiruan, pengujian ini menggunakan 100% atribut untuk mengetahui bagaimana hasil prediksi tanpa adanya *feature selection*.
3. Pengujian prediksi kelulusan mahasiswa menggunakan kombinasi algoritma jaringan syaraf tiruan dan information gain, pengujian ini melewati tahap pemilihan atribut dengan information gain lalu atribut terbaik akan dipakai algoritma jaringan syaraf tiruan untuk mengetahui hasil prediksi.
4. Hasil prediksi yang telah didapat dari pengujian algoritma jaringan syaraf tiruan dan hasil prediksi yang didapat dari pengujian kombinasi algoritma jaringan syaraf tiruan dengan information gain akan dianalisis hasilnya.

4.1 Hasil Pengujian Atribut Paling Berpengaruh dengan Information Gain

Berikut merupakan hasil dari proses *feature selection* [12] yang menghasilkan atribut paling berpengaruh beserta nilai dari information gain nya. Dataset yang dipakai adalah dataset mahasiswa tahun 2008. Proses *feature selection* ini mengurutkan atribut dengan nilai information gain dari besar ke kecil.

Tabel 3. Atribut Tertinggi Berdasarkan Nilai Information Gain (Tingkat 1)

Mata Kuliah	Nilai Information gain
Kalkulus 1	0.101765
Aljabar	0.100187
Kalkulus 2	0.090041
Prakinternet	0.074003
BING2	0.069880
BING1	0.066766
Prakfisika	0.063224
PTI	0.058520
Agama	0.055988

Bisnis	0.053687
--------	----------

Tabel 4. Atribut Tertinggi Berdasarkan Nilai Information Gain
(Tingkat 2)

Mata Kuliah	Nilai Information gain
Organisasi Arsitektur komputer	0.186169
Praktikum Basis Data	0.159710
Pemrograman Berorientasi Objek	0.155906
Probabilitas Dan Statistika	0.138762
Algoritma Struktur Data	0.137564
Rekaya Perangkat Lunak	0.131122
Logmat	0.122572
Bahasa Inggris III	0.104620
Prakom	0.102406
Agama	0.099178

Tabel 5. Atribut Tertinggi Berdasarkan Nilai Information Gain
(Tingkat 3)

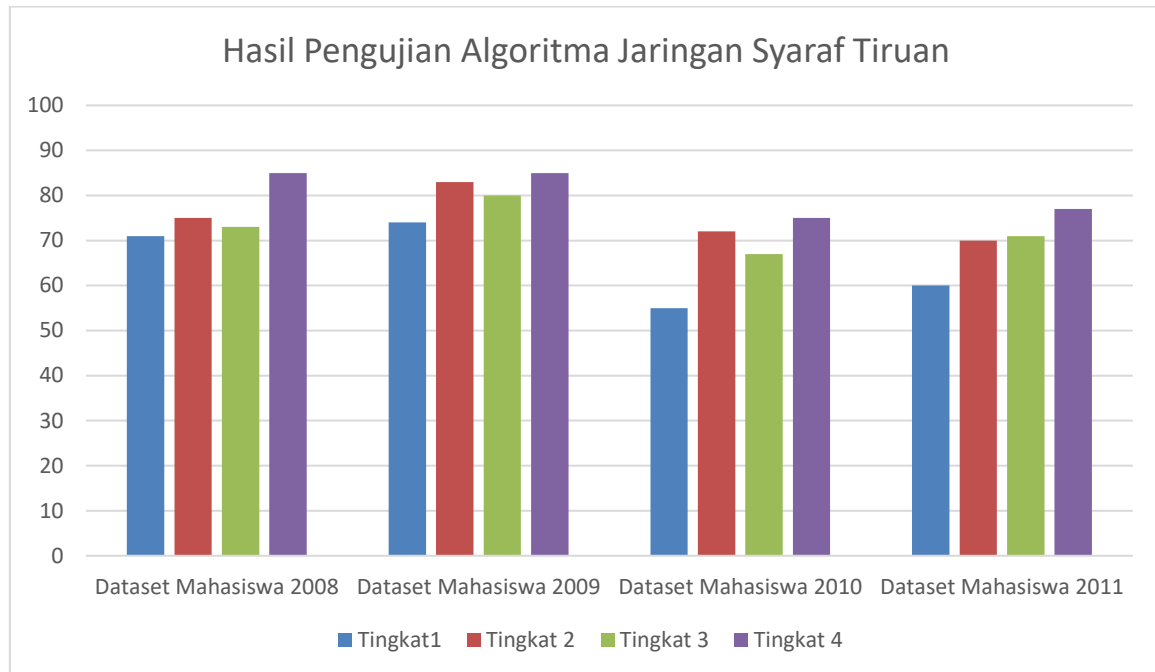
Mata Kuliah	Nilai Information gain
Sistem Operasi	0.174892
Teknik Berorientasi Objek	0.172810
Teori Komputasi	0.164790
Organisai Arsitektur Komputer	0.157366
Jaringan Komputer	0.133421
Kecerdasan Artifisial	0.125621
Probabilitas Dan Statistika	0.121946
Praktikum Basis Data	0.116454
DAA	0.113635
Pemrograman Berorientasi Objek	0.110583

Tabel 6. Atribut Tertinggi Berdasarkan Nilai Information Gain
(Tingkat 4)

Mata Kuliah	Nilai Information gain
Tugas Akhir 2	0.387251
Tugas Akhir 1	0.240045
Kecerdasan Artifisial	0.137086
Praktikum Basis Data	0.154211
Sistem Operasi	0.140564
Teori Komputasi	0.134143
Pemrograman Berorientasi Objek	0.132493
Kapita Selekt	0.131076
Algoritma Struktur Data	0.124619
Kewarganegaraan	0.116939

4.2 Pengujian Algoritma Jaringan Syaraf Tiruan

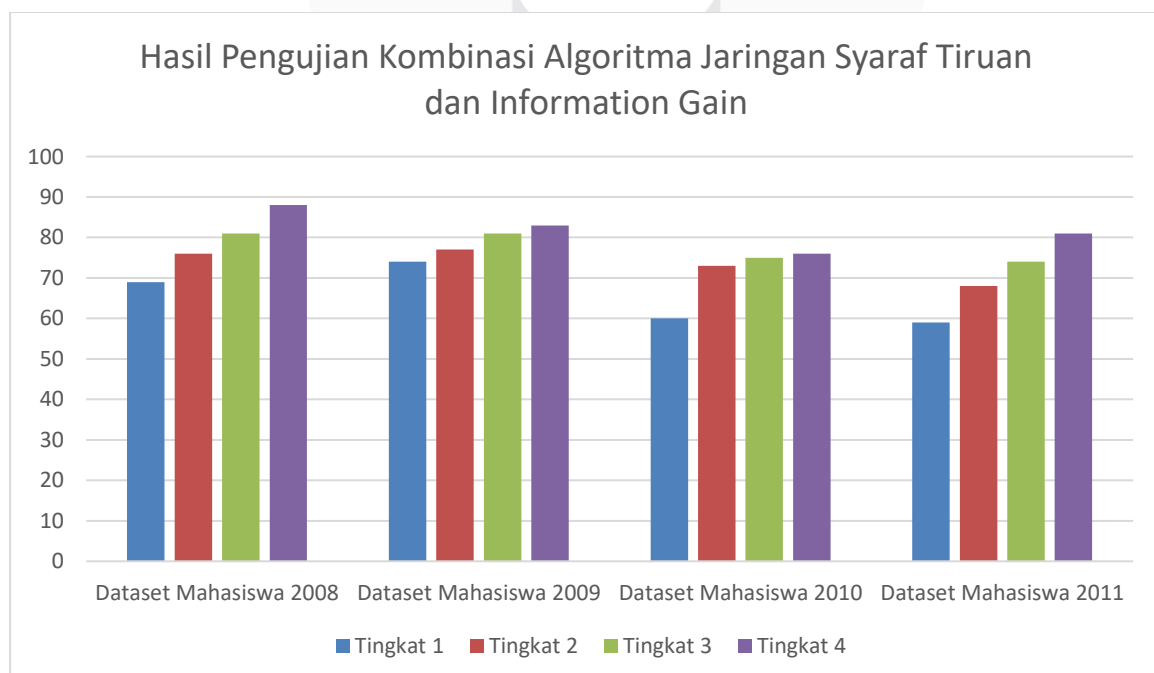
Pengujian ini dilakukan tanpa adanya pemilihan atribut, dan memakai keseluruhan atribut pada dataset mahasiswa S1 Teknik Informatika Universitas Telkom tahun 2008-2011. Berikut merupakan hasil akurasi yang didapat jika algoritma jaringan syaraf tiruan melakukan pengujian tanpa feature selection information gain.



Gambar 4. Hasil Pengujian Algoritma Jaringan Syaraf Tiruan

4.3 Pengujian Kombinasi Algoritma Jaringan Syaraf Tiruan dan Information Gain

Hasil pengujian utama dari penelitian ini adalah akurasi prediksi kelulusan mahasiswa S1 Teknik Informatika Universitas Telkom angkatan 2008-2011. Pengujian ini akan mengeluarkan hasil akurasi dari klasifikasi menggunakan algoritma jaringan syaraf tiruan dimana atribut yang dipakai adalah 60% dari total atribut dengan nilai information gain paling tinggi.



Gambar 5. Hasil Pengujian Kombinasi Algoritma Jaringan Syaraf Tiruan dan Information Gain

4.4 Analisis Hasil Pengujian

Berdasarkan hasil dari pemilihan atribut dengan information gain, mata kuliah yang berpengaruh ditingkat 1 merupakan mata kuliah berbasis perhitungan, dimana kalkulus 1 dan 2 lalu aljabar menempati 3 mata kuliah paling berpengaruh untuk tingkat 1. Lalu mata kuliah yang berpengaruh di tingkat 2 dan 3 merupakan mata kuliah yang sifatnya lanjutan dari mata kuliah dasar seperti sistem operasi, organisasi arsitektur komputer dan pemrograman berorientasi objek. Sedangkan mata kuliah yang berpengaruh di tingkat 4 adalah tugas akhir, mata kuliah ini sangat krusial karena tugas akhir merupakan syarat kelulusan bagi mahasiswa S1 Teknik Informatika Universitas Telkom

Berdasarkan pengujian menggunakan algoritma jaringan syaraf tiruan didapatkan hasil 55%-85% dengan penggunaan 100% atribut. Hasil yang didapat cukup variatif dengan penurunan akurasi ditingkat 3 ketinggian hampir disetiap tahun nya. Lalu pengujian menggunakan kombinasi algoritma jaringan syaraf tiruan dan information gain mendapat kenaikan hasil yang tidak begitu signifikan yaitu 59%-88%. Namun selalu terjadi peningkatan akurasi disetiap tingkatnya.

5. Kesimpulan

Kesimpulan dari pengujian ini adalah algoritma jaringan syaraf tiruan bisa dikombinasikan dengan information gain dan mendapat hasil yang sedikit lebih baik meskipun menggunakan 60% atribut yang tentunya lebih efisien daripada penggunaan 100% atribut. Berdasarkan hasil yang telah didapatkan, algoritma jaringan syaraf tiruan akan menghasilkan akurasi yang lebih baik jika dikombinasikan dengan information gain. Dengan hasil pengujian yang sudah dilakukan, diharapkan kelulusan mahasiswa di S1 Informatika Telkom University bisa dideteksi lebih awal sehingga aksi preventif bisa dilakukan oleh pihak mahasiswa dan program studi S1 Informatika guna meningkatkan tingkat kelulusan.

Saran untuk penelitian selanjutnya, semoga seluruh tahapan pengujian yang ada bisa dijadikan acuan untuk membuat model dengan akurasi yang lebih baik dan menggunakan data yang lebih variatif. Untuk penelitian selanjutnya nya direkomendasikan untuk melakukan pemilihan atribut dengan jumlah yang lebih optimal dan diharapkan adanya algoritma yang jauh lebih efektif dan efisien untuk deteksi awal kelulusan mahasiswa.

REFERENSI

- [1] Kusrini. 2009. Algoritma Data Mining. Yogyakarta : Penerbit ANDI
- [2] Han, Jiawei. 2000. Data Mining : Concepts and Techniques
- [3] Suhartono,Derwin.2012. "Dasar Pemahaman Neural Network"
- [4] Budiyanto,Utomo.2019. "Prediksi Kelulusan Tepat Waktu Mahasiswa Menggunakan Jaringan Syaraf Tiruan"
- [5] Prasetyo,Eko.2018. "Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree dan Artificial Neural Network"
- [6] Arunkumar,N.2018." Fully automatic model-based segmentation and classification approach for MRI brain tumor using artificial neural networks"
- [7] L. Z. Chen Xu, "Effectiveness Analysis of The Application of Clustering in Student," in International Conference on Education Technology and Information System (ICETIS 2013), 2013.
- [8] A. P. D. J. ., H. F. Fernandez, "E-learning and Educational Data Mining in Cloud Computing : An Overview," Int. J. Learning Technology, vol. 9, no. 1, pp. 25-52, 2014.
- [9] Pooja Kamavisdar, Sonam Saluja, et al., "A Survey on Image Classification Approaches and Techniques", International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, no. 1, 2013.
- [10] Risvik,Knut Magne. "Discretization of Numerical Attributes"
- [11] K,Saravanan. "Review on Classification Based on Artificial Neural Network"
- [12] sklearn. Mutual Information [Online] Available at: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mutual_info_score/ [Accessed 10 February 2021].