

BAB II

TINJAUAN PUSTAKA DAN LANDASAN TEORI

2.1 Tinjauan Pustaka

Penelitian yang pernah dilakukan oleh Cristóbal Romero, Sebastián Ventura, dan Enrique García dari Universitas Cordoba Jurusan *Computer Sciences and Numerical Analysis* pada tahun 2008 dengan judul *Data mining in course management systems: Moodle case study and tutorial*. Pada penelitian tersebut menjelaskan tentang proses penambangan data, di artikel tersebut menggunakan beberapa teknik yang ada pada *data mining* seperti statistik, visualisasi, klasifikasi, *clustering*, *association*, *pattern mining*, dan *text mining* (Romero, Ventura, dan García 2008).

Penelitian yang dilakukan oleh Rizqi Agung Permana dan Sucitra Sahara pada tahun 2009 dengan judul *Metode Support Vector Machine Sebagai Penentu Kelulusan Mahasiswa pada Pembelajaran Elektronik*. Pada artikel tersebut menjelaskan proses *data mining* dengan menggunakan metode *Support Vector Machine* (SVM). Atribut yang digunakan yaitu jenis kelamin, jumlah latihan, jumlah login, jumlah akses document, jumlah forum dan jumlah pesan. Hasil pengujian menggunakan metode tersebut didapatkan hasil akurasi sebesar 85.02%, dan nilai AUC (*Area Under Curve*) 0.610 (Permana dan Sahara 2019).

Pada penelitian yang berjudul *Data Mining Technique as Majors Support System Management with Classification Approach* yang diteliti oleh Muhammad Siddiq, Rosmawati Tamin, Sulihin Azis pada tahun 2019. Permasalahan yang diangkat adalah tes potensial akademik hanya digunakan sebagai standar dalam penerimaan mahasiswa baru, tetapi belum ada pengklasifikasian mahasiswa berdasarkan spesialisasi yang mereka punya. Pada penelitian tersebut dilakukan perbandingan dengan menerapkan metode-metode yang ada pada *data mining* untuk dianalisis. Hasil perbandingan dari SIAP-PPDB dengan *Major DSS*, dengan menggunakan metode *K-Nearest Neighbors Classification* (KNN) didapatkan tingkat akurasi sebesar 88,57% untuk *Major DSS* sedangkan SIAP-PPDB mendapatkan hasil akurasi 50%, dengan menggunakan metode *Classification Tree*

(CT) didapatkan hasil akurasi sebesar 97,14% untuk *Major DSS*, sedangkan untuk SIAP-PPBD mendapatkan 51,42% (Siddiq, Tamin, dan Azis 2019).

Penelitian yang berjudul Sistem Pendukung Keputusan Seleksi Asisten Laboratorium Dosen Elektro Menggunakan Metode Weighted Product Di Polines yang dilakukan oleh Satria Bayu Aji pada tahun 2015. Permasalahan yang menjadi objek penelitian yaitu sistem pendukung keputusan pemilihan calon asisten laboratorium di Politeknik Negeri Semarang. Pada penelitian tersebut menggunakan metode *Waighted Product* (WP), WP adalah suatu metode yang menghubungkan atribut dengan cara perkalian, kemudian bobot atribut yang sama nilainya harus dipangkatkan terlebih dahulu. Hasil yang didapat pada penelitian tersebut yaitu dapat menampilkan mahasiswa yang mendapatkan status rekomendasi untuk menjadi asisten dosen (Aji 2015).

Selanjutnya penelitian yang dilakukan oleh Kennedy, Dr. Purba Daru Kusuma M.T, dan Casi Setiamingsih S.T, M.T dengan judul penelitian *River Water Pollution Pattern Prediction Using A Simple Neural Network* pada tahun 2019, membahas tentang masalah pencemaran yang terjadi pada air sungai. Sebelumnya sudah banyak penelitian yang berkaitan untuk mengurangi pencemaran sungai. Pada penelitian tersebut menawarkan solusi pencemaran air sungai dengan sebuah sistem yang terdiri dari gabungan beberapa teknologi seperti *Sensor, Microcontroller, Internet of Things (IOT), Low-Power Wide-Area Network (LPWAN), Long Range (LORA)*, dan *Artificial Neural Network*. *Artificial Neural Network* diset didalam *microcontroller* yang berfungsi untuk melakukan pengolahan data. Hasil yang didapatkan adalah Jaringan Saraf Tiruan mampu menampilkan data praktis cukup akurat dengan hasil yang didapat akan ditampilkan didalam sebuah website (Kusuma dan Setianingsih 2019).

Penelitian yang dilakukan oleh Yuan Octavia D.P, A.N. Affandi, dan Hari Putranto pada tahun 2018 dengan penelitian yang berjudul *Studi Prakiraan Beban Listrik Menggunakan Metode Artificial Neural Network*. Pokok permasalahan yang menjadi fokus pada penelitian tersebut adalah kebutuhan energi listrik di Indonesia tiap tahunnya selalu meningkat, berdasarkan data 2011 sampai 2015

yang diambil dari Perusahaan Listrik Nasional (PLN). Sehingga penelitian tersebut dilakukan dengan tujuan yaitu melakukan perkiraan terhadap kebutuhan listrik kedepannya. Solusi yang dilakukan yaitu menerapkan *Artificial Neural Network* untuk dilakukan perkiraan beban listrik. Hasil yang didapat yaitu hasil *Mean Absolute Percentage Error* (MAPE) terbaik 0.512%, nilai *Learning Rate* yang didapat yaitu 0.9 dan untuk *Momentum* yaitu 0.1 (Octavia, Afandi, dan Putranto 2019).

2.2. Landasan Teori

2.2.1 Data Mining

Data Mining atau juga sering disebut *Knowledge Discovery in Database* (KDD). Merupakan suatu proses yang bertujuan untuk mendapatkan informasi dengan cara mengekstrak data yang ada di *database* menjadi informasi, kemudian informasi menjadi pengetahuan, dan pengetahuan menjadi suatu pengalaman. Pada proses *Knowledge Discovery in Database* terdiri dari gabungan beberapa teknik seperti Mesin Pembelajaran (*Machine Learning*), Statistik, pengenalan pola (*Pattern Recognition*) (Fayyad, Piatetsky-Shapiro, dan Smyth 1996). *Data Mining* secara umum bisa diterapkan terhadap semua jenis data selama data tersebut dapat digunakan. Pada umumnya data yang digunakan untuk *data mining* merupakan data yang berukuran besar seperti *database*, *data warehouse*, dan data transaksi ("*Data Mining: Concepts and Techniques 3rd Edition*," t.t.). *Data mining* juga bisa diterapkan untuk melakukan analisis data yang ada di suatu universitas untuk melihat peningkatan kemampuan mahasiswa berdasarkan *history* nilai yang ada (Ridwan, Suyono, dan Sarosa 2013).

Dalam melakukan pengolahan *data mining* ada 5 teknik pengolahan yang umum digunakan yaitu (Larose dan Larose 2014):

1. Deskripsi: Teknik deskripsi biasanya digunakan oleh para penulis dan analis untuk mencari pola yang ada didalam data.

2. Estimasi: Teknik estimasi digunakan para penulis untuk memperkirakan nilai dari *variabel* target dengan *variable* prediktor numerik.
3. Prediksi: Teknik prediksi mirip dengan teknik estimasi dan klasifikasi. Hanya saja sedikit perbedaan terletak di hasil akhir, hasil teknik prediksi ada dimasa depan.
4. Klasifikasi: Teknik klasifikasi mirip dengan estimasi. Perbedaanya terletak pada *variable* taget yang di kategorikan tidak berupa numerik.
5. *Clustering*: Teknik *clustering* merupakan kumpulan data yang mirip dengan yang lainnya, dan biasanya para penulis mengelompokkan data yang mirip kedalam kelas objek yang sama. Teknik *clustering* tidak sama dengan klasifikasi karena tidak adanya *variable* taget untuk *clustering*.

2.2.2 Asisten Dosen

Asisten Dosen adalah orang yang membantu dosen dalam melaksanakan tugas profesionalnya. Selain itu asisten dosen juga berperan sebagai penyambung lidah antara dosen dan mahasiswa. Menjadi seorang asisten dosen tentunya harus menguasai mata kuliah yang diajarkan oleh dosen, sehingga ketika dosen yang bersangkutan berhalangan untuk hadir maka asisten dosen harus siap untuk menggantikan mengajar. Oleh karna itu untuk mendapatkan seorang asisten dosen yang berkualitas maka dibutuhkan penyaringan dari kandidat-kandidat asisten dosen. Dengan harapan asisten dosen yang terpilih nantinya mampu membantu dosen dalam melaksanakan tugas profesionalnya dan mampu membantu mahasiswa dalam memahami pelajaran yang diajarkan.

2.2.3 Algoritma Neural Network

Algoritma *Neural Network* merupakan metode yang terinspirasi dari jaringan syaraf otak manusia karena didesain mengikuti cara otak manusia melakukan proses dan menyimpan suatu informasi (Hermanto 2010). Algoritma

Neural Network digunakan sebagai *tools* yang menggambarkan data statistik yang *non-linear*, dengan menggambarkan suatu hubungan yang kompleks antara *input* dan *output* (Singh dan Chauhan 2009).

Neural Network terdiri dari beberapa layer, disetiap layer biasanya terdapat minimal 1 atau lebih *Processing Elements* (PE). *Processing Elements* digunakan untuk mensimulasikan cara kerja *neuron* yang ada didalam otak manusia. Oleh karena itu PE juga sering disebut *neuron* atau *node*, setiap PE menerima input dari lapisan sebelumnya (Singh dan Chauhan 2009).

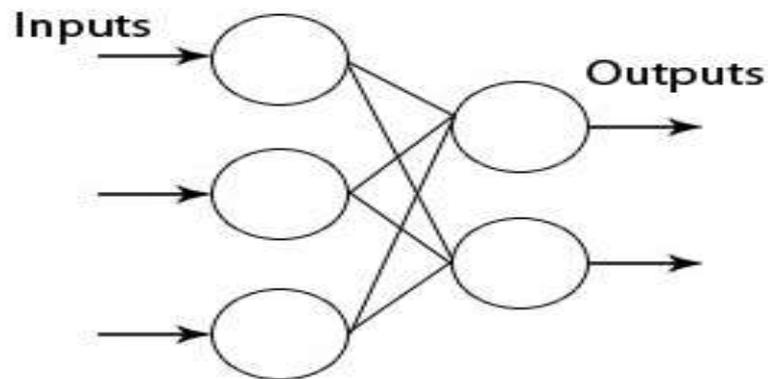
Menurut Rudolf Rudi Hermanto dalam penelitiannya yang berjudul *Neural Network dan Implementasinya Dalam Data Mining*, algoritma *neural network* memiliki 5 keunggulan yaitu (Hermanto 2010):

1. Akurasi yang tinggi.
2. Fleksibel dengan data yang *noise*.
3. Mudah dikelola dan dinamis.
4. Menggunakan pola yang paralel.
5. Independensi dari asumsi prior.

Neural Network terdiri dari neuron yang berhubungan antara satu neuron dengan neuron yang lainnya. Secara umum ada 3 jenis *Neural Network* yang sering digunakan, yaitu (Aprilla dkk. 2013):

1. *Single-Layer Neural Network*
2. *Multilayer Neural Network*
3. *Recurrent Neural Networks*

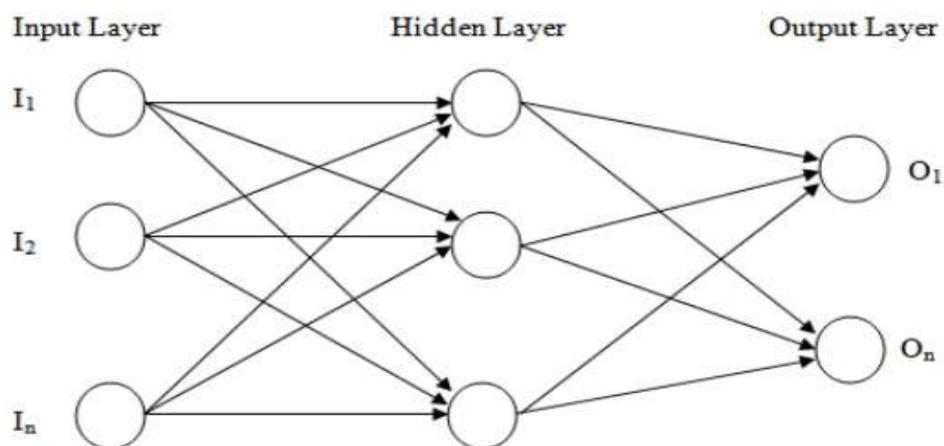
Single-Layer Neural Network hanya terdiri dari 2 layer yaitu *layer input* dan *layer output* seperti pada gambar 2.1. Setiap neuron yang ada di *input layer* berhubungan langsung ke neuron *output layer*. *Single-Layer Neural Network* hanya digunakan pada kasus yang tidak memiliki kompleksitas tinggi (Aprilla dkk. 2013).



Gambar 2. 1 Single-Layer Nural Network

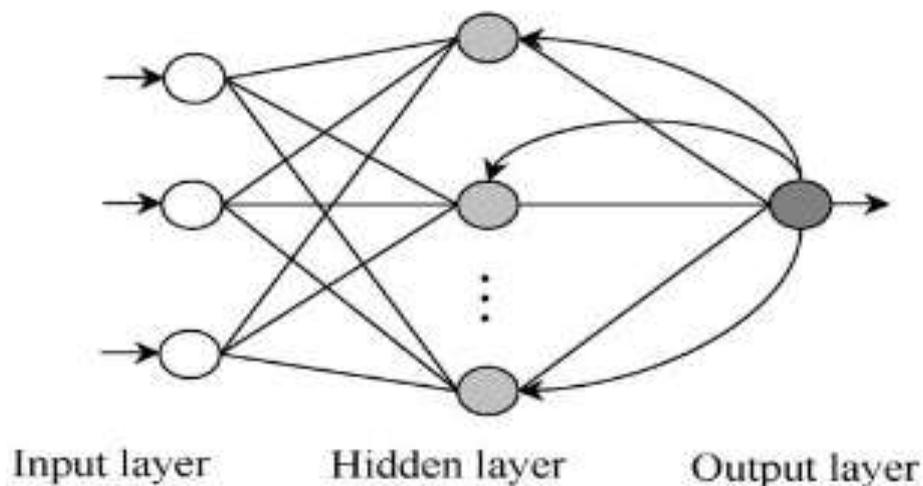
Multilayer Perceptron Neural Network memiliki satu *layer* tambahan yaitu *hidden layer* yang berada diantara *input layer* dan *output layer* seperti pada gambar 2.2. *Neural Network* jenis ini paling sedikit memiliki 3 *layer*, yaitu *input layer*, *hidden layer*, dan *output layer*. *Neural Network* jenis ini merupakan salah satu jenis jaringan yang sering digunakan (Aprilla dkk. 2013).

Multilayer Perceptron merupakan topologi *feedforward* yang mana informasi yang diterima hanya bergerak maju satu arah dari *input layer* melalui *hidden layer* dan *output layer* (Meinanda dkk. 2009). *Feedforward* merupakan topologi yang paling sederhana yang pernah dibuat sehingga topologi ini mudah untuk dipahami (Singh dan Chauhan 2009).



Gambar 2. 2 Multilayer Perceptron Nural Network

Recurrent Neural Networks hampir sama dengan *Multilayer Perceptron Neural Network*. Pada *Neural Network* jenis ini memiliki keunikan yaitu memiliki koneksi umpan balik dari *output layer* ke *layer* sebelumnya dapat dilihat pada gambar 2.3. Sehingga *Neural Network* jenis ini sering disebut dengan *backpropagation* karena pada jenis ini menggunakan topologi *backpropagation*, disebut *backpropagation* karena pada *Neural Network* jenis ini berlawanan dengan topologi *feedforward*. Hanya saja jenis *Neural Network* ini memiliki kelemahan karena adanya *loop* umpan balik dari *output layer* yang mengakibatkan terjadinya *time delay* (Aprilla dkk. 2013).



Gambar 2. 3 Recurrent Neural Networks

Didalam buku *Data Mining Concepts And Techniques* edisi ketiga karangan Jiawei Han, Micheline Kamber, dan Jian Pei telah diterangkan perhitungan rumus matematis *Neural Network* untuk menghitung *net input* ke unit, setiap *input* yang terhubung ke unit dikalikan dengan *weight* dan kemudian di jumlahkan dengan bias. Rumus perhitungan *Neural Network* sebagai berikut (*"Data Mining: Concepts and Techniques 3rd Edition," t.t.*):

$$I_j = \sum_i w_{ij} O_i + \theta_j$$

j = lapisan tersembunyi atau *output*

I_j = *Net Input*

w_{ij} = bobot koneksi dari unit i dilapisan sebelumnya ke unit j

O_i = *Output* dari unit i

θ_j = bias

2.2.4 *RapidMiner*

Software RapidMiner merupakan sebuah platform perangkat lunak yang bersifat *open source*, *RapidMiner* dibangun untuk melakukan analisis terhadap penambangan data, *text mining*, dan analisis prediksi (Hofmann dan Klinkenberg 2013). *Software RapidMiner* bekerja dengan menggunakan teknik deskriptif dan prediksi. *User RapidMiner* kini telah mencapai 500.000 lebih pengguna diseluruh dunia, sehingga tidak perlu diragukan lagi kemampuan *software* ini dalam melakukan penambangan data.

RapidMiner pertamakali dikembangkan pada tahun 2001 oleh 3 orang pengembang yaitu: RalfKlinkenberg, Ingo Mierswa, dan Simon Fischer di *Artificial Intelligence Unit* dari University of Dortmund, sebelum berubah nama menjadi *RapidMiner* sebelumnya bernama YALE (*Yet Another Learning Environment*) (Aprillia dkk. 2013)