

BAB II

KAJIAN PUSTAKA

2.1 Penelitian Relevan

Berdasarkan latar belakang yang sudah dijabarkan di atas, penelitian ini tidak lepas dari hasil penelitian sebelumnya, sebagai salah satu referensi tambahan dalam penelitian ini. Adapun penelitian yang telah di jadikan referensi tidak lepas dari pembahasan tentang sistem pendukung keputusan, Terdapat beberapa penelitian sebelumnya yang dapat dijadikan acuan dan perbandingan untuk melakukan penelitian ini , diantaranya :

Penelitian terkait pertama yang dilakukan oleh **Retnani Latifah, Emi Susilowati, Wulan Febriyanti** dengan judul **SISTEM PENDUKUNG KEPUTUSAN PENENTUAN CALON PENERIMA KARTU JAKARTA PINTAR (KJP) MENGGUNAKAN K-NEAREST NEIGHBOR**. Beberapa siswa di Jakarta memiliki kesempatan untuk mendapatkan kartu jakarta pintar (KJP) dari pemerintah provinsi DKI Jakarta. KJP merupakan program yang layak untuk diberikan bagi warga Jakarta dari kalangan masyarakat kurang mampu untuk mengambil pendidikan minimal sampai dengan tamat SMA/SMK dengan membiayai penuh oleh APBD Provinsi DKI Jakarta. Untuk melakukan pemilihan ini dengan memilih siswa mana saja yang berhak mendapatkan KJP. Seleksi dilakukan secara manual yaitu dengan cara melakukan pengecekan berkas-berkas setiap siswa. Hal ini sangat menyulitkan sekolah karena siswa yang dimiliki banyak Pemerintah provinsi Jakarta memiliki sebuah aplikasi web yang dapat digunakan sekolah untuk mendaftarkan siswa yang berhak menerima KJP. Namun, standar yang digunakan di aplikasi tersebut tidak sama dengan yang digunakan oleh beberapa sekolah sehingga menyulitkan dalam pengumpulan data diantaranya karena data siswa yang telah disetujui oleh sekolah ditolak oleh aplikasi. Penelitian ini bertujuan untuk merancang dan membangun suatu sistem pendukung keputusan yang sesuai dengan kebutuhan SMK XYZ dalam melakukan seleksi siswa yang berhak mendapatkan KJP. Proses penentuan siswa akan dipilih menggunakan metode *K-Nearest Neighbor* (KNN) yang merupakan metode populer untuk melakukan klasifikasi. Dari hasil penelitian yang telah dilakukan dapat diketahui bahwa KNN bekerja dengan cukup baik dalam memberikan rekomendasi siswa mana yang dapat diajukan untuk menjadi penerima KJP. Hasil akurasi tertinggi diperoleh dengan nilai 91,67% dengan $k=3$ pada hasil pembagian fold keempat. Rata-rata akurasi tertinggi dari 5 cross fold adalah 74.93% dengan menggunakan $k=3$ (Retnani, Susilowati, & Febriyanti, 2018). Penelitian terkait ke dua yang dilakukan oleh **Faris Asrory, Joseph Dedy Irawan, Abdul Wahid** dengan judul **APLIKASI PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU MENGGUNAKAN METODE KNN**. Perguruan tinggi adalah Perguruan tinggi adalah tahap akhir opsional pada pendidikan. Al Mahsuni dapat

dikatakan salah satu perguruan tinggi populer jika di bandingkan perguruan tinggi lain. Namun STITNU Al Mahsuni memiliki pandangan dan duta besar untuk menjadi Universitas besar dan di perhitungkan sebagai Perguruan tinggi berkualitas. Untuk mencapai tujuan tersebut, di perlukan izin dari bawah naungan pemerintah. agar mendapatkan akreditasi, lembaga yang bersangkutan harus memenuhi syarat yang kuat. salah satunya yang terpenting ialah lulus tepat waktu. haruslah berbanding lurus. namun pada kenyataannya masih masih ada peserta didik yang yang lulus pada semester yang lebih dari waktu yang di tentukan di STITNU Al Mahsuni. Dari permasalahan tersebut dibutuhkan sebuah sistem yang mampu menangkap petunjuk keterlambatan kelulusan mahasiswa sejak dini. Metode yang digunakana adalah KNN. Algoritma *K-Nearest Neighbor* adalah metode yang menggunakan pengelompokan ketetanggaan (neighbor) sebagai nilai prediksi dari query instance yang baru. Hasil nilai akurasi tertinggi adalah prediksi dengan menggunakan k 1,4,5,6,8,9 dan 10 yang masing-masing memiliki nilai akurasi 93,33%. Selanjutnya nilai akurasi tertinggi ke 2 adalah prediksi dengan nilai k 3 dan 7 yang masingmasing memiliki nilai akurasi 91,11%. Sedangkan prediksi dengan nilai k 2 memilki nilai akurasi 84,44%, yang juga sekaligus dapat di kategorikan sebagai nilai akurasi paling rendah. Kesimpulannya adalahPrediksi menggunakan metode k-NN dengan parameter jenis kelamin, memiliki pekerjaan/membiayai kuliah secara mandiri, status pernikahan, umur dan jurusan SMA pada mahasiswa STINU Al Mahsuni Lombok Timur dapat dilakukan (Asrory, Irwan, & Wahid, Vol. 4 No. 2, September 2020). Penelitian terkait ke tiga yang di lakukan oleh Muhammad Kafil dengan judul **PENERAPAN METODE K-NEAREST NEIGHBORS UNTUK PREDIKSI PENJUALAN BERBASIS WEB PADA BOUTIQ DEALOVE BONDOWOSO**. Toko Dealove Bondowoso merupakan toko besar yang menyediakan berbagai macam alat kebutuhan yang berhubungan dengan pakaian. Sistem penjualan yang ada di toko dialove masih dilakukan secara acak dalam pencatatan barang yang keluar atau terjual. Sehingga masih banyak memiliki berbagai kendala yang dihadapi. masalah yang dihadapi yaitu sulit dalam mendata barang yang keluar, penjumlahan kedaan barang yang tidak baik, serta proses penjumlahan harga yang sering terjadi kesalahan, perekapan yang terjadi bisa dibilang kurang baik karena dikerjakan secara manual. Permintaan konsumen yang tidak menentu seiring berjalannya waktu, mengakibatkan kekurangan stok di gudang. Tujuan dari pada di adakannya peramalan kebijakan adalah untuk memperoleh informasi data mengenai perubahan di masa mendatang yang akan berpengaruh terhadap implementasi kebijakan serta pemikirannya. Dari permasalahan tersebut, disarankan menggunakan metode K-NN untuk memprediksi penjualan pada toko Dealove dengan masa periode setiap bulan dari Bulan Januari 2016 – September 2018. Berdasarkan penerapan dan pengujian yang dilakukan dapat ditarik kesimpulan

bahwa hasil pengujian keakuratan metode menggunakan menggunakan 12 data training dan 12 data testing diperoleh hasil nilai akurasi 83,3% dan nilai error sebesar 16,7 % ooooo (kafil, 2019).

Dari paparan perbandingan ke tiga jurnal penelitian di atas dengan penelitian selanjutnya adalah didapatkan persamaan menggunakan metode yang sama tapi data yang digunakan berbeda. Sehingga mendapatkan informasi yang berbeda.

2.2 Landasan Teori

Landasan teori menjadi suatu pondasi dan gambaran terkait teori-teori yang ada pada penelitian ini yang tersusun secara terstruktur sehingga penelitian ini berjalan dengan sistematis. Berikut beberapa landasan teori yang digunakan pada penelitian ini adalah. Klasifikasi, sistem pendukung keputusan, bantuan siswa miskin, metode KNN, python.

2.2.1 Klasifikasi

Klasifikasi adalah Proses pengkategorian atau proses pembelajaran yang diawasi digunakan untuk mengatur data secara sistematis. Klasifikasi ini merupakan pendekatan untuk mendiagnosis penyakit jantung menggunakan algoritma pengelompokan. dan untuk meningkatkan klasifikasi pengklasifikasian sebagai pelengkap dari pengklasifikasi (Lia, Sukemi, & Dian).

2.2.2 Data Mining

Data mining adalah penambahan data yang menggabungkan pembelajaran mesin, pengenalan pola, statistik, basis data, dan teknologi visualisasi untuk memecahkan masalah pengambilan informasi dari basis data besar. Data Mining adalah analisis otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya.

Tahap-tahap data mining adalah sebagai berikut:

- Pembersihan Data(Data Cleaning) pembersih data merupakan proses menghilangkan kegaduhan dan data yang tidak tepat waktu
- Data utuh (Data Integration) integrasi data merupakan proses penggabungan data dari berbagai data besar ke dalam satu data besar baru.
- Pemilihan Data data yang ada pada data besar sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari data besar.

- Presentasi Pengetahuan(Knowledge presentation) merupakan pencitraan dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh user.

2.2.3 Bantuan Siswa Miskin

Bantuan Siswa Miskin (BSM) adalah kebijakan pembangunan pendidikan pemerintah yang ditujukan untuk meningkatkan akses pendidikan dasar dan menengah yang berkualitas bagi peserta didik dari rumah tangga miskin/miskin. Keberhasilan program bantuan ini seharusnya meningkatkan kualitas pencapaian) yang menjadi indikator keberhasilan program wajib belajar pendidikan dua belas tahun (Suginam, Nasution, Lubis, & Mesran, 2018).

2.2.4 Metode KNN

KNN adalah merupakan metode menyusun data yang mengelompokkan data baru berdasarkan jarak antara data baru dengan data tertentu/tetangga terdekat. Terdapat training data set di KNN yang salah satunya berisi nilai dari class, dan prediktor, yang nantinya akan di gunakan sebagai salah satu syarat untuk menentukan kategori yang sesuai berdasarkan prediktor, dan data set uji memiliki konten data baru klasifikasi akan di lakukan sesuai dengan model yang akan dibuat kemudian, dan akurasi klasifikasi aka dievaluasi. Teknologi KNN sendiri memiliki langkah-langkah sebagai berikut yaitu dimulai dari input. Data training, label data training, k, data testing (Sahambangung, Munir, & Sanger, 2018).

- Pada keseluruhan data uji, menghitung jaraknya ke setiap data training
- Menentukan K data uji coba yang jaraknya paling dekat dengan data. Dalam penentuan nilai K, nominal yang akan dipilih harus berangka ganjil yang semakin kecil nilai yang dipilih akan menjadi lebih akurat nilai yang diperoleh.
- Pengujian
- Memeriksa identitas dari data k ke data ini
- Menentukan identitas yang frekuensinya paling banyak memasukan data uji ke kelas dengan frekuensi paling banyak
- Berhenti

Langkah-langkah cara perhitungan KNN berdasarkan Euclidean Distance :

- Menentukan parameter k. Nilai k ditentukan oleh setengah atau lebih dari data

- Tentukan jarak terdekat dengan barisan K
- Cocokkan kelas yang tepat
- Temukan jumlah kelas ini dari tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelas data yang akan dievaluasi.

Hitungan jarak antara data yang akan dianalisis dengan semua data pelatihan atau data sampel. Untuk menghitung jaraknya dengan menggunakan rumus Euclidean Distance. Perhitungan jarak antara data baru dengan data lama dilakukan menggunakan rumus sebagai berikut.

$$: d_i = \sum_{i=1}^n (p_1 - q_1)^2$$

Keterangan x1 = sampel data

X2 = data uji data testing

2.2.5 Python

Python adalah bahasa pemrograman yang serbaguna dan mudah dimengerti python juga dapat berjalan di berbagai rencana kerja sistem operasi, seperti halnya windows (python), menggunakan python untuk menghasilkan matriks mode tunggal Samsul Arifin, Indra Bayu Muktyas. 2 Linux, Mac OS, Android (Goole Play), dan lain lain. Dengan menggunakan program Python, dapat dibangkitkan suatu matriks unimodular sehingga dapat menghasilkan suatu SPL dengan solusi tunggal semua bilangan bulat. Dalam tulisan ini akan dikaji mengenai suatu matriks unimodular yang dapat dijadikan sebagai matriks koefisien dalam suatu SPL, (Arifin & Muktyas, 2018)