

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Aplikasi sistem pendukung keputusan untuk menentukan penjurusan pada tingkat SMA sudah beberapa kali dilakukan dengan menggunakan metode yang bermacam-macam. Selain menggunakan metode *K-Nearest Neighbor* (KNN), aplikasi penjurusan ini dibuat dengan metode *fuzzy* (*fuzzy decision making method*).

Metode *fuzzy* yang digunakan pada aplikasi penjurusan ini adalah *Weighted Subsethood Based Rule Generation Algorithm* (WSBA) untuk pembuatan *fuzzy rules*nya. Penerapan *fuzzy* WSBA ini sesuai untuk penentuan bobot pada kasus pengelompokan data. Selain itu *rules* yang dihasilkan lebih akurat. Pengujian sistem ini meliputi pengujian fungsionalitas dan pengujian hasil keluaran sistem dengan cara membandingkan keluaran sistem dengan prestasi siswa setelah dijuruskan. Hasil pengujian keluaran sistem ini menunjukkan adanya ketidaksesuaian hasil sebesar 8,5% untuk jurusan IPA dan 9,4% untuk jurusan IPS (Komalasari, 2005).

Kedua metode yang digunakan untuk sistem penjurusan ini sama-sama baik, namun pada metode *fuzzy* masih terdapat ketidaksesuaian hasil. Sedangkan menggunakan metode *K-Nearest Neighbor* terdapat beberapa kelebihan, selain penghitungannya yang sederhana, proses trainingnya cepat (Nugroho, 2006).

2.2 Sistem Penjurusan

Penjurusan atau course yang ditawarkan di level pendidikan menengah telah diterapkan di Indonesia sejak jaman Belanda. Sekolah HBS yang merupakan sekolah tingkat menengah untuk anak-anak Eropa, dan AMS yang merupakan sekolah menengah atas untuk anak-anak pribumi. Pertama kalinya jurusan pada sekolah AMS ini terbagi menjadi 2 jurusan yaitu kelompok A (bahasa) dan kelompok B (*science*). Namun, pada masa-masa selanjutnya penjurusan di Indonesia diterapkan sejak memasuki SMP, yang kemudian dihapus pada tahun 1962. Sistem penjurusan kemudian hanya dikenal di SMA dengan 3 macam jurusan, yaitu A(*sains*), B(bahasa/budaya), dan C(*social*). Pengistilahan ini mengalami perubahan dan spesifikasi pada masa-masa berikutnya, seperti A1, A2, A3, A4. Dan akhirnya kembali lagi seperti sekarang, penjurusan tidak lagi menggunakan huruf. Melainkan menggunakan kategori, yaitu IPA, IPS, dan Bahasa.

Penjurusan pada tingkat SMA ini diperkenalkan dengan tujuan supaya para siswa dapat mengenali minat dan bakat mereka sesuai dengan kemampuan akademiknya. Siswa-siswi yang memiliki kemampuan eksakta yang baik, biasanya diarahkan untuk memilih jurusan IPA. Sedangkan siswa yang memiliki minat di bidang sosial dan ekonomi akan diarahkan untuk memilih jurusan IPS. Dan untuk siswa yang gemar berbahasa maka akan diarahkan untuk memilih jurusan Bahasa.

Pengarahan sejak dini dimaksudkan untuk memudahkan para siswa dalam memilih/menentukan minor/bidang ilmu

yang akan ditekuninya di Universitas atau akademi. Hal ini tentunya juga akan mengarah pula kepada karirnya kelak. Tetapi penjurusan di tingkat SMA tidak menjamin bahwa seorang siswa akan memilih bidang studi yang sama di Universitas. Karena pada kenyataannya, banyak siswa program IPA memilih jurusan Ekonomi, Ilmu Komunikasi, Hubungan Internasional, dll. Begitu juga dengan siswa jurusan IPS memilih jurusan di bidang science seperti Teknik Informatika, Teknik Arsitektur, Teknik Industri, dll. Pemilihan jurusan yang berbeda dengan bidang yang ditekuni semasa SMA adalah hal yang cukup wajar. Hal ini dikarenakan anak seusia SMA belum dapat memastikan karirnya (Ramli, 2008).

Saat ini, hal penting yang perlu dipikirkan adalah apakah penjurusan di SMA sudah efektif, terutama jika dipandang dari sudut pandang siswanya. Pada Undang-Undang Sisdiknas 2003 tentang tujuan pendidikan menengah, terdapat 2 arahan yaitu mempersiapkan siswa ke jenjang perguruan tinggi, dan untuk terjun ke masyarakat (bekerja). Penjurusan pada tingkat SMA dapat dikatakan efektif apabila ditangani dengan sungguh-sungguh.

2.3 Sistem Pendukung Keputusan

Sistem pendukung keputusan (dalam istilah Inggrisnya: *Decision Support System* atau sering disingkat **DSS**) merupakan bagian dari sistem informasi berbasis komputer yang dipakai untuk mendukung pengambilan keputusan pada suatu organisasi atau perusahaan. Dapat juga dikatakan

sebagai sistem komputer yang mengolah data menjadi informasi untuk mengambil keputusan dari masalah semi-terstruktur yang spesifik. Beberapa pengertian sistem pendukung keputusan menurut para ahli komputer, Little (1970) mendefinisikan sistem pendukung keputusan sebagai sebuah himpunan/kumpulan prosedur berbasis model untuk memproses data dan pertimbangan guna membantu manajemen dalam mengambil suatu keputusan. Menurut Keen (1980) sistem pendukung keputusan adalah sistem berbasis komputer yang dibangun lewat sebuah proses adaptif dari pembelajaran, pola-pola penggunaan dan evolusi sistem. Bonczek (1980) mendefinisikan sistem pendukung keputusan sebagai sebuah sistem berbasis komputer yang terdiri atas beberapa komponen antara lain komponen sistem bahasa (*language*), komponen sistem pengetahuan (*knowledge*), komponen sistem pemrosesan masalah (*problem processing*) yang saling berinteraksi satu dengan lainnya. Hick (1993) menyebutkan sistem pendukung keputusan sebagai sekumpulan *tools* komputer yang terintegrasi untuk memungkinkan seorang *decision maker* untuk berinteraksi langsung dengan komputer guna menciptakan informasi yang berguna dalam membuat keputusan semi terstruktur dan tidak terstruktur yang tidak terantisipasi (Republik BM). Sedangkan menurut Alter (1990), apabila sistem pendukung keputusan dibandingkan dengan sebuah sistem pemrosesan data elektronik (PDE / *Electronic Data Processing*) maka akan terlihat 5 hal perbedaaan yaitu (Subakti, 2002):

Tabel 2.1 Tabel Perbandingan SPK dengan PDE

Pembanding	SPK	PDE
Penggunaan	Aktif	Pasif
Pengguna	Manajemen	Operator / Pegawai
Tujuan	Efektifitas	Efisiensi Mekanis
Time Horizon	Sekarang dan masa depan	Masa lalu
Kelebihan	Fleksibilitas	Konsistensi

Berdasarkan definisi di atas, dapat ditarik beberapa ciri / karakteristik umum dari sebuah sistem pendukung keputusan yang ideal yaitu :

- a. SPK adalah sebuah sistem berbasis komputer dengan antarmuka antara mesin/komputer dan pengguna.
- b. SPK ditujukan untuk membantu membuat keputusan dalam menyelesaikan suatu masalah dalam berbagai level manajemen dan bukan untuk mengganti posisi manusia sebagai pembuat keputusan.
- c. SPK mampu memberi alternatif solusi bagi masalah semi/tidak terstruktur baik bagi perseorangan ataupun kelompok dan dalam berbagai macam proses pengambilan keputusan.
- d. SPK menggunakan data, basis data dan analisa model-model keputusan.
- e. SPK bersifat adaptif, efektif, interaktif, *easy to use* dan *fleksibel*.

Sistem pendukung keputusan terbangun dari beberapa subsistem, antara lain (Subakti, 2002):

a. Subsistem Manajemen Data

Subsistem ini meliputi basisdata yang relevan dengan keadaan yang ada, serta dikelola oleh sebuah system yang dikenal sebagai *database management system* (DBMS).

b. Subsistem Manajemen Model

Subsistem ini merupakan sebuah paket perangkat lunak yang berisi model-model *financial*, *statistic*, *management science* dan model kuantitatif lain yang menyediakan kemampuan analisis sistem dan manajemen perangkat lunak yang terkait.

c. Subsistem Manajemen Pengetahuan (Knowledge)

Subsistem ini merupakan subsistem yang mampu mendukung subsistem yang lain atau berlaku sebagai sebuah komponen yang mampu berdiri sendiri.

d. Subsistem Antarmuka Pengguna (User Interface)

Subsistem ini merupakan media tempat komunikasi antara pengguna dan system pendukung keputusan serta tempat pengguna untuk memberikan perintah kepada system pendukung keputusan.

Beberapa tahapan untuk pembuatan sistem pendukung keputusan, antara lain:

a. Pendefinisian masalah

b. Pengumpulan data atau elemen informasi yang relevan.

c. Pengolahan data menjadi suatu informasi baik dalam bentuk laporan grafik maupun tulisan.

- d. Menentukan alternatif-alternatif solusi (bisa dalam bentuk persentase).

Sedangkan tujuan dari sistem pendukung keputusan itu sendiri, antara lain:

- a. Membantu menyelesaikan masalah.
- b. Mendukung manajer dalam mengambil suatu keputusan.
- c. Meningkatkan efektifitas bukan efisiensi pengambilan keputusan.

Dapat terlihat jelas, bahwa sistem pendukung keputusan mempunyai beberapa manfaat penting yaitu:

- a. Membantu pengambilan keputusan yang rasional, sesuai dengan jenis keputusan yang diperlukan.
- b. Dapat membuat peramalan/*forecasting*.
- c. Membandingkan alternatif tindakan.
- d. Membuat analisis dampak.
- e. Membuat model

2.4 Data Mining

Data mining merupakan proses mengekstraksi pola-pola yang menarik (tidak sepele, implisit, tak-diketahui sebelumnya, mungkin bermanfaat) dari data yang berukuran besar (*Hand, 2001*). Ada beberapa istilah yang mempunyai kemiripan dengan data mining, yaitu ekstraksi pengetahuan, analisis pola, dan pengerukan data. Ada yang berpendapat juga bahwa *data mining* merupakan sinonim dari istilah *knowledge discovery in database (KDD)*.

Data mining muncul berdasarkan fakta bahwa pertumbuhan data yang sangat pesat, tetapi miskin akan pengetahuan. Alasan memilih data mining dibanding

analisis data secara tradisional, adalah (Widyastuti, 2008):

- a. Data mining mampu menangani jumlah data kecil sampai data yang berukuran terabyte.
- b. Mampu menangani data yang mempunyai banyak dimensi, yaitu puluhan sampai ribuan dimensi.
- c. Mampu menangani data dengan kompleksitas yang tinggi.

Data Mining memiliki 2 macam pendekatan, yaitu :

- a. Pendekatan Deskriptif, yaitu pendekatan dengan cara mendeskripsikan data masukan. Metode yang termasuk pada pendekatan ini, adalah :
 - a.1 Metode Deskripsi konsep / kelas, yaitu data dapat diasosiasikan dengan kelas atau konsep.
 - a.2 Metode Aturan Asosiasi, metode yang membuat aturan berdasarkan kondisi yang sering terjadi.
- b. Pendekatan Prediktif, yaitu pendekatan yang dapat digunakan untuk memprediksi, dengan hasil berupa kelas atau cluster. Metode yang termasuk pada pendekatan ini adalah :
 - b.1 Metode Klasifikasi dan Prediksi, data yang digunakan untuk membentuk model yang mendeskripsikan kelas data yang penting, atau model yang memprediksikan trend data.
 - b.2 Metode *Clustering* , yaitu mengelompokkan data untuk membentuk kelas-kelas baru atau sering disebut *cluster*.

2.5 Metode *K-Nearest Neighbor* (KNN)

Algoritma *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap obyek berdasarkan data pembelajaran yang jaraknya paling dekat dengan obyek tersebut. Prinsip kerja dari *K-Nearest Neighbor* (KNN) adalah mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga (neighbor) terdekatnya dalam data pelatihan.

Pada fase pembelajaran, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi dari data pembelajaran. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk data test (yang klasifikasinya tidak diketahui). Jarak dari vektor yang baru ini terhadap seluruh vektor data pembelajaran dihitung, dan sejumlah k buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut.

Nilai k yang terbaik untuk algoritma ini tergantung pada data. Secara umum, nilai k yang tinggi akan mengurangi efek *noise* pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur. Nilai k yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan *cross-validation*. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain, $k = 1$) disebut algoritma *k-nearest neighbor*.

Ketepatan algoritma k -NN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan

relevansinya terhadap klasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur agar performa klasifikasi menjadi lebih baik.

Sesuai dengan prinsip kerja K-Nearest Neighbor yaitu mencari jarak terdekat antara data yang akan dievaluasi dengan k tetangga (*neighbor*) terdekatnya dalam data pelatihan. Persamaan dibawah ini menunjukkan rumus perhitungan untuk mencari jarak terdekat dengan d adalah jarak dan p adalah dimensi data (Agusta, 2007):

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \quad (1)$$

Dengan keterangan :

x_1	: sampel data	d	: jarak
x_2	: data uji	p	: dimensi data
i	: variable data		