

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Sekolah Menengah Atas (SMA), adalah jenjang pendidikan menengah di Indonesia setelah lulus Sekolah Menengah Pertama (SMP) atau sederajat. SMA ditempuh dalam kurun waktu 3 tahun, mulai dari kelas X sampai kelas XII[1]. Dalam kurun waktu 3 tahun tersebut, setiap jurusan pada SMA terdapat materi pelajaran yang sama salah satunya yaitu pelajaran matematika. Matematika merupakan materi pelajaran yang dipadatkan atau ditambahkan karena disesuaikan dengan materi pelajaran Standar Internasional sehingga pendidikan di dalam negeri dan pendidikan di luar negeri dapat seimbang.

Dalam sebuah penelitian yang berhubungan dengan data, dibutuhkan sebuah metode yang dapat membantu dalam proses pelaksanaannya. Data mining merupakan disiplin ilmu yang mempelajari tentang metode untuk mengekstrak pengetahuan atau pola dari suatu data, sehingga data mining juga sering disebut *knowledge discovery in database*[2]. Data mining dapat digunakan untuk pengelompokkan data, prediksi, estimasi, dan menentukan kaidah asosiasi dalam suatu data yang ada. Perlunya data mining karena adanya sejumlah data besar yang dapat digunakan untuk menghasilkan informasi dan *knowledge* yang berguna. Informasi dan *knowledge* yang didapat tersebut dapat digunakan untuk mengetahui suatu pola dalam suatu data yang banyak, terlebih lagi besarnya kebutuhan untuk mengubah data tersebut menjadi informasi yang berguna. Metode data mining bermacam-macam disesuaikan dengan kebutuhan yang ada[3].

Salah satu penelitian yang menggunakan data mining adalah estimasi dengan menggunakan model regresi linear. Model regresi adalah istilah lain untuk jenis model linear[2]. Analisis regresi adalah metodologi populer yang mengeksplorasi hubungan antara variabel respon dan satu atau lebih variabel

penjelas. Metodologi ini telah banyak diterapkan dibanyak bidang seperti ekonomi, teknik, biologi, bisnis, dan sebagainya[4]. Selain bidang tersebut, model regresi juga dapat digunakan pada bidang pendidikan, contohnya untuk menghitung estimasi dalam sebuah perhitungan khususnya untuk estimasi nilai siswa. Dalam sejumlah data hubungan sebenarnya jarang dapat diketahui akan tetapi hubungan tersebut dapat diestimasi berdasarkan data pengamatan. Metode yang populer digunakan adalah estimasi kuadrat terkecil. Hal ini didasarkan pada kenyataan bahwa ia meminimumkan jumlah kuadrat perbedaan nilai yang diharapkan dengan nilai observasi[5].

Belum begitu banyak penelitian yang membahas tentang estimasi nilai mata pelajaran khususnya nilai SMA. Adanya estimasi nilai bagi pihak sekolah berfungsi sebagai acuan dalam proses belajar mengajar karena dengan menerapkan estimasi pada setiap nilai mata pelajaran, dapat menjadi tolak ukur dan membantu dalam proses meningkatkan kualitas belajar apabila nilai yang diperoleh tidak memenuhi syarat ketuntasan. Estimasi dalam sebuah penelitian ini membahas tentang nilai matematika yang berfungsi untuk mengetahui hubungan antara variabel nilai Ulangan Harian Terprogram (UHT) 1 dan nilai Ulangan Harian Terprogram (UHT) 2 terhadap variabel nilai mid.

Salah satu penelitian yang pernah dilakukan oleh Yuliana[5] mengenai penerapan model regresi linear robust dengan estimasi M pada data nilai kalkulus II mahasiswa Universitas Widya Dharma Klaten. Dalam penelitian tersebut, yang dimaksudkan subjek penelitiannya, yaitu mahasiswa angkatan 2012/2013. Dari subjek penelitian, diambil data nilai kalkulus I dan nilai kalkulus II diambil saat subjek kuliah berada pada semester II. Metode penelitian yang digunakan penulis dalam penelitian tersebut meliputi studi pustaka dan studi kasus. Data yang digunakan pada penelitian ini merupakan data nilai ujian akhir semester II mata kuliah kalkulus II sebagai variabel dependen dan nilai ujian akhir semester I mata kuliah kalkulus I dan trigonometri sebagai variabel independen.

Berdasarkan pertimbangan di atas, penulis akan melakukan penelitian dengan mengambil objek dari SMA Kesatrian 1 Semarang yang beralamat di Jalan Pamularsih No. 116 Semarang. Di SMA Kesatrian 1 Semarang terdapat 3 jurusan, yaitu IPA, IPS, dan Bahasa. Masing-masing jurusan tersebut terdapat mata pelajaran matematika yang siswa-siswi pelajari dari kelas X sampai kelas XII. Dengan menggunakan variabel nilai matematika, yaitu nilai Ulangan Harian Terprogram (UHT) 1 dan nilai Ulangan Harian Terprogram (UHT) 2 akan diketahui pengaruh kedua nilai tersebut terhadap nilai mid.

Dalam penelitian ini, penulis menggunakan pendekatan data mining dengan penerapan model regresi linear berganda akan dilakukan untuk estimasi nilai siswa untuk mengetahui pengaruh nilai ulangan harian terhadap nilai mid dengan menggunakan algoritma regresi linear. Penelitian ini menggunakan RMSE (*Root Mean Square Error*) untuk menghitung standar error hasil estimasi. Diharapkan dengan perhitungan estimasi tersebut dapat diperkirakan hasil nilai yang kemungkinan didapat pada saat ujian mid sehingga pihak SMA dapat meningkatkan kualitas belajar siswanya apabila nilai kurang memenuhi syarat ketuntasan. Berdasarkan penelitian-penelitian sebelumnya, metode atau model yang baik digunakan untuk menghitung estimasi adalah model regresi linear. Dengan menggunakan Matlab berguna untuk mengetahui hasil dari kinerja algoritma yang digunakan.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, diperoleh rumusan masalah, yaitu bagaimana menerapkan model regresi linear berganda untuk mengestimasi nilai mata pelajaran matematika siswa semester I dari SMA Kesatrian 1 Semarang sehingga dapat diketahui hubungan antara nilai UHT 1 dan UHT 2 terhadap nilai mid.

1.3 Batasan Masalah

Agar penyusunan tugas akhir ini tidak terlalu meluas dan keluar dari pokok permasalahan yang dirumuskan, maka penulis perlu memberikan batasan masalah sebagai berikut:

1. Dataset yang diambil untuk penelitian ini difokuskan pada siswa semester I dari SMA Kesatrian 1 Semarang tahun ajaran 2015/2016.
2. Dataset yang penulis analisa ini difokuskan untuk estimasi nilai matematika dengan parameter yang digunakan meliputi nilai UHT 1, nilai UHT 2, dan nilai mid.
3. Algoritma yang digunakan untuk mengestimasi nilai tersebut adalah algoritma regresi linear dengan model regresi linear berganda dan menggunakan perhitungan hasil estimasi standar error RMSE (*Root Mean Square Error*).
4. Penggunaan Matlab sebagai piranti perangkat lunak yang digunakan untuk melihat hasil kinerja dari algoritma yang digunakan terhadap dataset.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang diuraikan di atas, maka tujuan dari penelitian ini adalah mengestimasi nilai matematika SMA Kesatrian 1 Semarang pada semester I untuk mengetahui hubungan antara variabel nilai UHT 1 dan nilai UHT 2 terhadap nilai mid pada masing-masing jurusan (IPA, IPS, dan Bahasa) dengan menggunakan algoritma regresi linear berganda.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini diantaranya adalah:

- a. Manfaat bagi Akademik

Penelitian ini dapat memberikan informasi bagi peneliti atau calon peneliti lain untuk menerapkannya ke dalam sistem yang lebih luas dan lebih kompleks atau sebagai bahan acuan yang dapat dikembangkan.

- b. Manfaat bagi Siswa SMA Kesatrian 1 Semarang

Dapat membantu siswa dalam tolak ukur nilai yang akan didapat pada saat mid dengan mengacu pada nilai ulangan harian sehingga siswa yang nilai hariannya kurang memenuhi syarat ketuntasan dapat meningkatkan belajarnya.

c. Manfaat bagi SMA Kesatrian 1 Semarang

Membantu pihak sekolah untuk meningkatkan kualitas belajar siswanya apabila nilai kurang memenuhi syarat ketuntasan.

d. Manfaat bagi Universitas Dian Nuswantoro

Mengetahui sejauh mana kemampuan mahasiswa dalam menguasai materi yang telah didapat selama masa perkuliahan di Universitas Dian Nuswantoro.

e. Manfaat bagi penulis

Sebagai sarana untuk menerapkan dan mengembangkan ilmu yang telah didapatkan selama perkuliahan serta menambah pemahaman dan pengalaman terutama tentang mata kuliah data mining.