**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNOLOGI INFORMASI**

**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**

**USULAN TUGAS AKHIR**

# IDENTITAS PENGUSUL

**NAMA : M. Rizal Syah Karim**

**NRP : 5110100138**

**DOSEN WALI : Daniel O. Siahaan, S.Kom., M.Sc., PDEng**

**DOSEN PEMBIMBING : 1. Dr. Ir. R.V. Hari Ginardi, M.Sc.**

**2. Dr. Chastine Fatichah, M.Kom.**

# JUDUL TUGAS AKHIR

“Prediksi Nilai Mata Kuliah berdasarkan Nilai Prasyarat Menggunakan Metode *Support Vector Regression*”

# LATAR BELAKANG

*Educational Data Mining (EDM)* merupakan hal yang cukup penting di dunia pendidikan. Tujuan utama dari EDM adalah untuk membangun model komputasional serta *tools* nya dalam menggali data yang terdapat dalam dunia pendidikan. Penggalian data di dunia pendidikan *(EDM)* merupakan hal penting karena penyelenggara pendidikan bisa mendapatkan informasi terkait dari data yang berjumlah besar dan beragam. Oleh karena itu diperlukan metode yang baik bukan hanya dalam mengolah data tetapi juga dalam memperoleh informasi yang tersimpan dalam data tersebut.

Salah satu aplikasi dari penggalian data dalam dunia pendidikan ini adalah mencari variabel yang berhubungan dengan nilai akademik. Nilai merupakan parameter dari penilaian sukses atau tidaknya siswa. Oleh karena itu, dengan diketahuinya variabel yang berhubungan dengan nilai akademik siswa dapat membantu penyelenggara pendidikan ataupun siswa dalam melakukan tindakan ke depan yang sesuai.

Didalam tingkat level pendidikan yang lebih tinggi yakni tingkat universitas, mata kuliah prasyarat acapkali diisi dengan materi dasar dimana materi tersebut akan digunakan dalam mata kuliah semester-semester selanjutnya. Matakuliah prasyarat adalah matakuliah yang merupakan persyaratan untuk suatu matakuliah yang diprasyarati. Apabila suatu matakuliah mempunyai matakuliah prasyarat tertentu, maka pengambilannya hanya dibenarkan setelah persyaratannya dipenuhi[[1](#FMI13)]. Dari sini penulis tertarik mencari seberapa signifikan nilai yang didapat pada mata kuliah prasyarat berpengaruh terhadap nilai mata kuliah selanjutnya.

Ada berbagai macam metode yang dapat digunakan untuk mencari hubungan antar variabel. Metode yang umumnya digunakan dalam menentukan korelasi antara variabel adalah regresi linier[[2](#MBy08)]. Hanya saja permasalahannya, regresi yang biasa digunakan hanya untuk menentukan *one-to-one correlation*[[3](#Agr97)]. Untuk menentukan korelasi yang bersifat *non-linier* dibutuhkan metode regresi yang kompleks[[4](#Jui12)]. Pada tugas akhir ini, dipilih metode *Support Vector Regression.* Metode ini dikenal sebagai metode klasifikasi yang handal dan juga dapat digunakan untuk memecahkan masalah regresi.

# RUMUSAN MASALAH

Rumusan masalah yang diangkat dalam Tugas Akhir ini adalah sebagai berikut:

1. Bagaimana penerapan metode *Support Vector Regression* untuk memprediksi Nilai Mata Kuliah berdasarkan Nilai Mata Kuliah Prasyarat*?*
2. Bagaimana hubungan antara nilai mata kuliah prasyarat dengan nilai mata kuliah yang diprasyarati dengan menggunakan metode statistik?

# BATASAN MASALAH

Permasalahan yang dibahas dalam tugas akhir ini memiliki beberapa batasan, yaitu sebagai berikut :

1. Dataset yang dipakai berupa nilai mata kuliah nilai prasyarat dan mata kuliah nilai yang diprasyarati yang didapat dari sim akademik Teknik Informatika ITS dalam 5 tahun terakhir.
2. Metode yang digunakan untuk memprediksi adalah *Support Vector Regression.*
3. Bahasa pemrograman yang digunakan adalah *python*.

# TUJUAN PEMBUATAN TUGAS AKHIR

Tugas akhir ini bertujuan untuk :

1. Menerapkan metode *Support Vector Regression* untuk memprediksi Nilai Mata Kuliah berdasarkan Nilai Mata Kuliah Prasyarat.
2. Analisa hubungan antara nilai mata kuliah prasyarat dengan nilai mata kuliah yang diprasyarati.

# MANFAAT TUGAS AKHIR

Tugas akhir ini dikerjakan dengan harapan dapat memberikan manfaat yang besar untuk dunia pendidikan, khususnya teknik informatika its. Dengan diketahuinya hubungan antara nilai syarat dengan peluang keberhasilan studi, maka pada saat sebelum dosen yang mengajar mata kuliah lanjutan, dosen yang bersangkutan akan mendapatkan *early warning* dari sistem, mahasiswa mana yang butuh bimbingan lebih lanjut dan mana yang sudah menguasai materi yang diberikan.

# TINJAUAN PUSTAKA

1. **Mata Kuliah Prasyarat**

Matakuliah prasyarat adalah matakuliah yang merupakan persyaratan untuk suatu matakuliah yang diprasyarati. Apabila suatu matakuliah mempunyai matakuliah prasyarat tertentu, maka pengambilannya hanya dibenarkan setelah persyaratannya dipenuhi. Dengan demikian apabila seorang mahasiswa membatalkan suatu matakuliah prasyarat, semua matakuliah yang diprasyarati juga dinyatakan batal[[1](#FMI13)].

Gambar 1. Contoh Pohon Prasyarat Mata Kuliah Teknik Informatika ITS

**Gambar 1**. Contoh Pohon Prasyarat Mata Kuliah Teknik Informatika ITS

Dalam jurusan teknik informatika ITS, terdapat beberapa mata kuliah yang bisa diambil apabila prasyaratnya terpenuhi. Dalam Gambar 1, dicontohkan bahwa untuk dapat mengambil mata kuliah analisis data multivariat, mahasiswa harus lulus terlebih dahulu mata kuliah statistika komputasional. Untuk dapat mengambil mata kuliah statistika komputasional, maka diwajibkan lulus terlebih dahulu mata kuliah aljabar linier.

1. **Support Vector Machine**

*Support Vector Machine* (SVM) adalah algoritma klasifikasi yang berbasis pada statistical learning theory dan merupakan metode yang bersifat supervised. Pada proses pelatihannya, SVM mencari sebuah fungsi hyperplane sebagai pemisah antara dua kelas yang paling optimal, yaitu yang memiliki jarak terjauh dari kedua kelas tersebut.

Misalkan terdapat L buah data dengan D buah atribut (berdimensi D) yang tergolong pada salah satu kelas yi = -1 atau +1, sehingga memiliki bentuk yang dapat dituliskan sepertipada persamaan 1.

(1)

Untuk data tersebut, diasumsikan bahwa semua data pada kedua kelas dapat dipisahkan dengan sebuah hyperplane secara linier yang memiliki bentuk yang dapat dituliskan seperti persamaan 2.

(2)

sehingga semua data pada kedua kelas memiliki bentuk yang dapat dituliskan seperti persamaan 3 dan persamaan 4.

(3)

(4)

dan dapat dituliskan seperti pada persamaan 5.:

(5)

Jarak terdekat antara hyperplane dengan koordinat (0,0) ditulis pada persamaan 6.

(6)

Tujuan pelatihan SVM adalah memaksimalkan jarak tersebut dengan meminimalkan ||w||, sehingga dapat ditulis sebagai persamaan 7.

(7)

Untuk data yang tidak dapat dipisahkan secara linier, SVM mentransformasikan data pada input space tersebut ke *feature space* yang berdimensi lebih tinggi menggunakan fungsi kernel. Data pada ruang yang baru tersebut kemudian dapat dipisahkan secara linier.

1. **Regresi Linier**

Regresi Linier merupakan metode prediksi dalam penggalian data. Regresi Linier memprediksi nilai variabel respon yang didasarkan dari satu atau lebih variabel prediktor. Pada dasarnya, prediksi pada regresi dimodelkan dalam persamaan aljabar matematika yang dirumuskan pada persamaan 8.

 (8)

Dimana Y merupakan varibel respon yang diprediksi, x merupakan varibel prediktor, sedangkan α dan β merupakan koefisien pada persamaan hasil analisis regresi.

1. **Support Vector Regression**

*Support Vector Regression* (SVR) adalah pengembangan dari SVM untuk melakukan regresi untuk memprediksi nilai keluaran dari suatu masukan. Regresi dapat dilakukan secara linier dan nonlinier. Seperti SVM, SVR juga awalnya dikembangkan untuk membentuk fungsi linier. Untuk melakukan regresi nonlinier, data masukan ditransformasikan ke feature space yang berdimensi lebih tinggi yang berkorelasi dengan nilai keluaran. Tujuan SVR adalah mencari fungsi regresi seperti yang ditulis pada persamaan 9.

(9)

yang memiliki deviasi maksimal sebesar ɛ dari nilai target pada semua data training sehingga masalah optimasi yang harus diselesaikan adalah yang ditulis pada persamaan 10.

(10)

# RINGKASAN ISI TUGAS AKHIR

Blok diagram sistem secara umum dipaparkan pada Gambar 2.



Gambar . Blok Diagram Sistem

## Pengumpulan Data

Data masukan pada Tugas Akhir ini adalah kumpulan data yang berisi nilai mahasiswa yang diperoleh melalui sim akademik Teknik Informatika ITS. Nilai tersebut terdiri dari nilai-nilai mata kuliah prasyarat sebagai variabel prediktor dan nilai mata kuliah yang diprasyarati sebagai kelas atribut yang akan diprediksi atau variabel respon.

Dataset yang digunakan dapat dilihat padaTabel 1.

Tabel 1.Dataset Nilai Mahasiswa

|  |  |  |  |
| --- | --- | --- | --- |
| No | Nama Dataset | Deskripsi | Jumlah *record* |
| 1 | dataset1 | Nilai mata kuliah prasyarat 1tahun 2009-2013 | ±500 |
| 2 | dataset2 | Nilai mata kuliah prasyarat 2 tahun 2009-2013 | ±500 |
| 3 | dataset3 | Nilai mata kuliahyang diprasyarati tahun2009-2013 | ±200 |

## Preprocessing Data

Data yang didapatkan kemudian di preprocessing.sebelum dilakukan training, dataset akan dipreprocessing terlebih dahulu dengan menggunakan data cleansing dan transformasi data. Preposesing dilakukan untuk meningkatkan kualitas data dan hasil prediksi.

## Pelatihan Data menggunakan metode Support Vector Regression

Tahap ini merupakan tahap pembelajaran yang dilakukan oleh sistem. Pada tahap ini dibentuk model prediksi dari data latih yang akan digunakan pada tahap uji.Pada tahap ini classifier yang digunakan adalah SVR. Classifier yang telah dilatih digunakan untuk melakukan prediksi nilai mata kuliah yang diprasyarati.

## Evaluasi dan Pengujian Data

Tahap uji dilakukan untuk melakukan prediksi dari data uji menggunakan model yang telah dibuat pada tahap latih.Classifier yang telah dilatih digunakan untuk melakukan prediksi nilai mata kuliah yang diprasyarati.

## Penghitungan Akurasi dan Error

tahap ini berguna untuk menguji model yang dibuat pada tahap latih. Dari hasil prediksi yang ada dapat dihitung kesalahan hasil prediksi jika diketahui nilai sebenarnya. Nilai mata kuliah yang diprasyarati hasil prediksi dibandingkan dengan nilai mata kuliah yang diprasyarati sebenarnya. Akurasi hasil prediksi ditampilkan dalam *Root Mean Square Error (RMSE)* dan *Normalized Mean Square Error (NMSE)*yang dituliskan pada persamaan 11, 12. Metrik-metrik tersebut mengukur kedekatan antara nilai prediksi dengan nilai sebenarnya. Semakin kecil nilai-nilai tersebut, maka semakin tinggi akurasi prediksi.

(11)

(12)

# METODOLOGI

## Penyusunan proposal tugas akhir

Penyusunan proposal tugas akhir merupakan tahap awal dalam proses pengerjaan tugas akhir. Dalam proposal ini diajukan metode *Support Vector Regression* untuk memprediksi nilai mata kuliah berdasarkan nilai prasyarat .

## Studi literatur

Literatur yang akan dipelajari antara lain metode *Support Vector Machine, Linier Regression, Support Vector Regression*, *Support Vector Regression* untuk prediksi nilai siswa dan *Preceeding Educational Data Mining.*

## Implementasi perangkat lunak

Perangkat lunak yang dibuat dikhususkan untuk membuktikan akurasi metode *Support Vector Regression*. Dengan menggunakan library *libsvm*[[5](#Cha11)] dan implementasi program menggunakan bahasa phyton dengan IDE portable phyton dan matlab.

## Pengujian dan evaluasi

Pada tugas akhir ini classifier yang digunakan adalah SVR. Classifier yang telah dilatih digunakan untuk melakukan prediksi nilai mata kuliah yang diprasyarati. Akurasi hasil prediksi ditampilkan dalam Root *Mean Square Error (RMSE)* dan *Normalized Mean Square Error (NMSE).* Metrik-metrik tersebut mengukur kedekatan antara nilai prediksi dengan nilai sebenarnya. Semakin kecil nilai-nilai tersebut, maka semakin tinggi akurasi prediksi. Semakin tinggi akurasi prediksi, semakin tinggi korelasi antar variable.

## Penyusunan Buku Tugas Akhir

Pada tahap ini dilakukan penyusunan laporan yang menjelaskan dasar teori dan metode yang digunakan dalam tugas akhir ini serta hasil dari implementasi aplikasi perangkat lunak yang telah dibuat. Sistematika penulisan buku tugas akhir secara garis besar antara lain:

1. Pendahuluan
   1. Latar Belakang
   2. Rumusan Masalah
   3. Batasan Tugas Akhir
   4. Tujuan
   5. Metodologi
   6. Sistematika Penulisan
2. Tinjauan Pustaka
3. Desain dan Implementasi
4. Pengujian dan Evaluasi
5. Kesimpulan dan Saran
6. Daftar Pustaka

# JADWAL KEGIATAN

Pengerjaan tugas akhir ini akan dilakukan mengikuti rencana pengerjaan seperti yang ditunjukkan oleh Tabel 2.

.

Tabel 2. Jadwal Pengerjaan Tugas Akhir

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Tahapan | 2013 | | | | | | | | | | | | | | | | | 2014 | | | |
| September | | | | Oktober | | | | November | | | | Desember | | | | | Januari | | | |
| Penyusunan Proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Studi Literatur |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Perancangan sistem |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Implementasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pengujian dan evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Penyusunan buku |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

# DAFTAR PUSTAKA

x

|  |  |
| --- | --- |
| [1] | M Byrne and B Flood, "Examining the relationship among background variable and academic performance of first year accounting students at an Irish University," *Journal of Accounting Education*, vol. 26, no. 4, pp. 202-212, December 2008. |
| [2] | A Agresti and B Finlay, "Statistical methods for the social sciences," 1997. |
| [3] | Fu Jui-Hsi and Chang Jui-Hung, "A Support Vector Regression-based Prediction of Student School Performance," 2012. |
| [4] | FMIPA UGM. (2013) Ketentuan Ketentuan Akademik. [Online]. <http://mipa.ugm.ac.id/web/content/ketentuan-ketentuan-akademik> |
| [5] | Chang C and Lin C J, "LIBSVM : a library for support vector machines," *ACM Transaction on Intelligent System and Technology*, 2011. |

x

.