

Tinjauan Mata Kuliah

Data Mining secara umum dapat diartikan sebagai proses penambangan data yang menghasilkan sebuah *output* (keluaran) berupa pengetahuan. Perkembangan yang cepat dalam teknologi pengumpulan dan penyimpanan data telah memudahkan manusia untuk mengumpulkan sejumlah data berukuran besar sehingga menghasilkan gunung data. Ekstraksi informasi yang berguna dari basis data tersebut menjadi pekerjaan yang cukup menantang. Seringkali alat dan teknik analisis data tradisional tidak dapat digunakan dalam mengekstrak informasi dari data berukuran besar. Teknik data mining digunakan untuk memeriksa basis data berukuran besar sebagai cara untuk menemukan pola yang baru dan berguna.

Data mining memiliki dua peran utama, yaitu: 1) prediktif untuk memprediksi nilai dari atribut tertentu berdasarkan pada nilai dari atribut-atribut lain dan 2) deskriptif untuk menurunkan pola-pola (korelasi, *trend*, *cluster*, trayektori, dan anomali) yang meringkas hubungan yang pokok dalam data. Secara khusus, data mining menggunakan ide-ide seperti (1) pengambilan contoh, estimasi, dan pengujian hipotesis dari statistika dan (2) algoritma pencarian, teknik pemodelan, dan teori pembelajaran dari kecerdasan buatan, pengenalan pola, dan *machine learning*. Komponen dalam sebuah arsitektur data mining terdiri dari basis data, basis pengetahuan, mesin data mining, modul evaluasi pola, dan antar muka grafis. Tahapan proses data mining sesuai dengan CRISP-DM terdiri dari: Fase Pemahaman Data, Fase Persiapan Data, Fase Pemodelan, Fase Evaluasi, dan Fase Penyebaran. Sedangkan pemanfaatan data mining dapat dilakukan pada berbagai bidang, misalkan pada bidang pendidikan, kesehatan, perbankan, riset, rekayasa manufaktur, sampai dengan bioinformatika.

Setelah mempelajari Mata Kuliah MSIM4403 Data Mining (3 SKS), mahasiswa dapat menerapkan konsep dan teknik-teknik dalam data mining dengan tepat. Mahasiswa akan belajar mengaplikasikan teknik estimasi, teknik klasifikasi, teknik analisis asosiasi dan klasterisasi, serta mengevaluasi teknik dan menerapkannya pada sebuah aplikasi/tools. Pada bahan ajar ini juga didukung penggunaan aplikasi WEKA sebagai aplikasi yang sesuai untuk implementasi teknik-teknik dalam data mining.

Capaian Pembelajaran Utama dari mata kuliah ini adalah mahasiswa diharapkan mampu menerapkan konsep dan teknik-teknik data mining dalam berbagai bidang. Sedangkan secara khusus, setelah mempelajari mata kuliah mata kuliah ini mahasiswa diharapkan mampu untuk:

1. menjelaskan konsep dasar data mining;
2. mendefinisikan pemodelan dan persiapan data;
3. menjelaskan konsep teknik-teknik dalam data mining;
4. membedakan teknik estimasi dalam data mining;
5. membedakan teknik klasifikasi dalam data mining;
6. membedakan teknik analisis asosiasi dalam data mining;
7. membedakan teknik analisis klasterisasi dalam data mining;
8. menghitung kinerja teknik-teknik dalam data mining;
9. menggunakan aplikasi data mining dalam berbagai bidang.

Sesuai dengan beban SKS-nya, bahan ajar MSIM4403/Data Mining ini dikemas dalam 9 modul yang terdiri dari berikut ini.

- Modul 1. Konsep Dasar Data Mining, membahas tentang pengertian data mining, kebutuhan data mining, ilmu yang berkaitan dengan data mining, dan teknik basis data.
- Modul 2. Pemodelan dan Persiapan Data, membahas tentang pengertian data, jenis data dan kualitas data, data *preprocessing*, dan visualisasi data.
- Modul 3. Konsep Teknik-teknik dalam Data Mining, membahas tentang konsep dasar *predictive* data mining dan konsep dasar *descriptive* data mining.
- Modul 4. Teknik Estimasi dalam Data Mining, membahas tentang teknik regresi linier sederhana dan regresi linier berganda.
- Modul 5. Teknik Klasifikasi dalam Data Mining, membahas tentang algoritma ID3, C4.5, *Nearest Neighbor* dan *Bayesian*.
- Modul 6. Teknik Analisis Asosiasi dalam Data Mining, membahas tentang algoritma Apriori dan *FP Growth*.
- Modul 7. Teknik Analisis Klasterisasi dalam Data Mining, membahas tentang algoritma *K-Means* dan *Hierarchical Agglomerative Clustering (HAC)*.
- Modul 8. Evaluasi Kinerja Teknik dalam Data Mining, membahas tentang kinerja algoritma *predictive* data mining dan kinerja algoritma *descriptive* data mining.
- Modul 9. Data Mining dengan Aplikasi WEKA, membahas tentang pengenalan dan prapemprosesan data, prediksi data dan deskripsi data dengan WEKA.

Agar tujuan mata kuliah data mining ini dapat di kuasai, sebaiknya pelajari modul secara berurutan sesuai dengan susunan modul yang terdapat pada bahan ajar MSIM4403. Pelajari modul secara lengkap mulai dari penyajian, kemudian dilanjutkan dengan mengerjakan latihan, dan mengerjakan tes formatif untuk mengukur penguasaan terhadap materi modul ini. Terakhir cocokkan jawaban latihan dan tes formatif yang disediakan, jika masih ada yang belum cocok, tuntaskan sampai semuanya cocok.

Dengan begitu diharapkan akan dapat menguasai materi modul pemrograman berbasis *desktop* ini dengan lebih baik.

Selamat belajar dan selamat berlatih, semoga sukses mengikuti mata kuliah ini.

PETA KOMPETENSI
MSIM4403/Data Mining/3 sks

