**Penerapan Data MiningUntuk Evaluasi Kinerja**

**Akademik Mahasiswa Menggunakan**

**Algoritma *Naive Bayes Classifier***



**Oleh :**

**Orien Patriana**

**09031181520035**

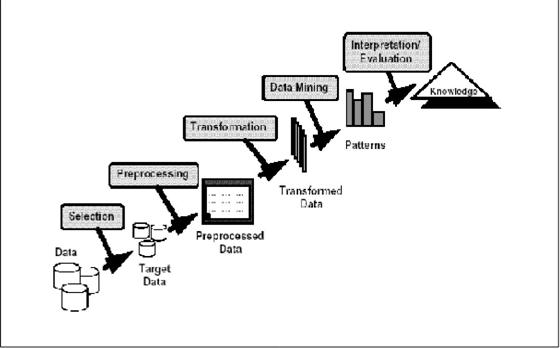
**SISTEM INFORMASI**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS SRIWIJAYA**

**2018**

**Knowladge Discovery In Database** **(KDD)** adalah proses menemukan pengetahuan yang berguna dari kumpulan data. Teknik penambangan data yang banyak digunakan ini adalah proses yang meliputi persiapan dan pemilihan data, pembersihan data, menggabungkan pengetahuan sebelumnya pada set data dan menafsirkan solusi akurat dari hasil yang diamati. Area aplikasi KDD utama termasuk pemasaran, deteksi penipuan, telekomunikasi dan manufaktur.

[](http://2.bp.blogspot.com/-AAonayVJI5E/UWjdj02drzI/AAAAAAAAAok/yRa8-0tac44/s1600/4d.jpg)

**Tahapan Proses KDD**

**1. Data Selection**

* Menciptakan himpunan data target , pemilihan himpunan data, atau memfokuskan pada subset variabel atau sampel data, dimana penemuan (discovery) akan dilakukan.
* Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

**2. Pre-processing/ Cleaning**

* Pemprosesan pendahuluan dan pembersihan data merupakan operasi dasar seperti penghapusan noise dilakukan.
* Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus KDD.
* Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi).
* Dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

**3. Transformation**

* Pencarian fitur-fitur yang berguna untuk mempresentasikan data bergantung kepada goal yang ingin dicapai.
* Merupakan proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses ini merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data

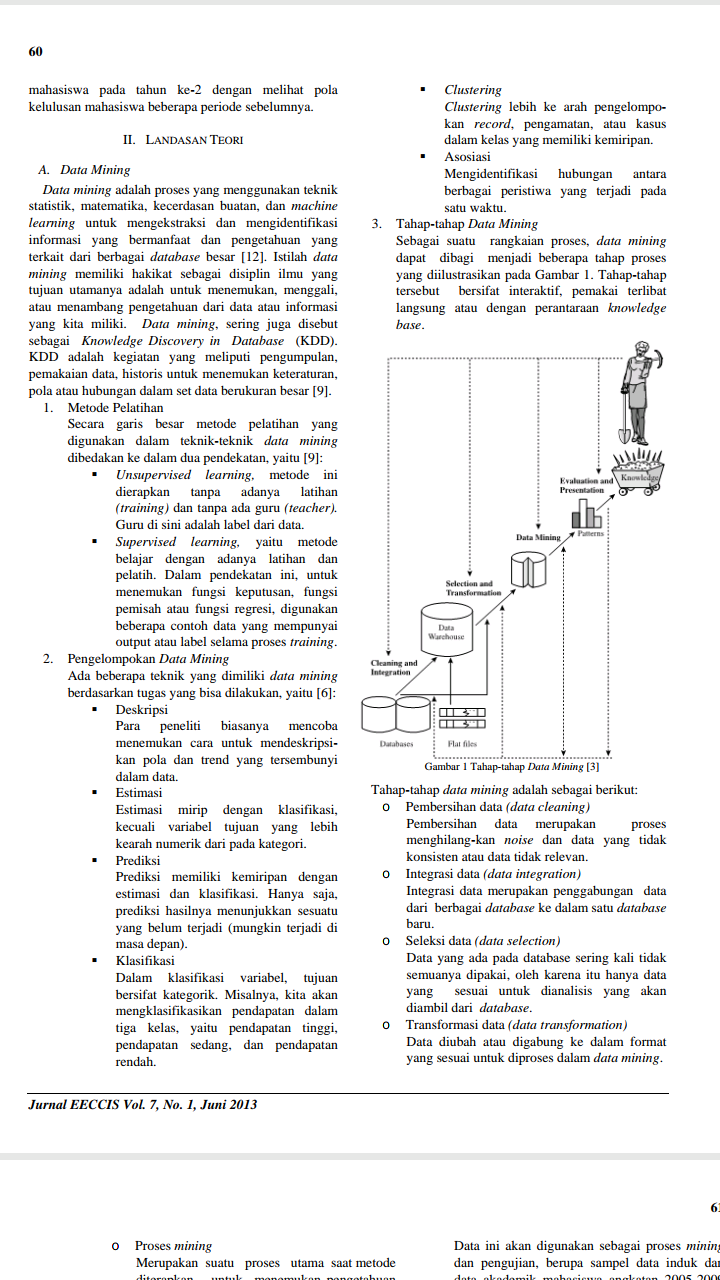
**4. Data mining**

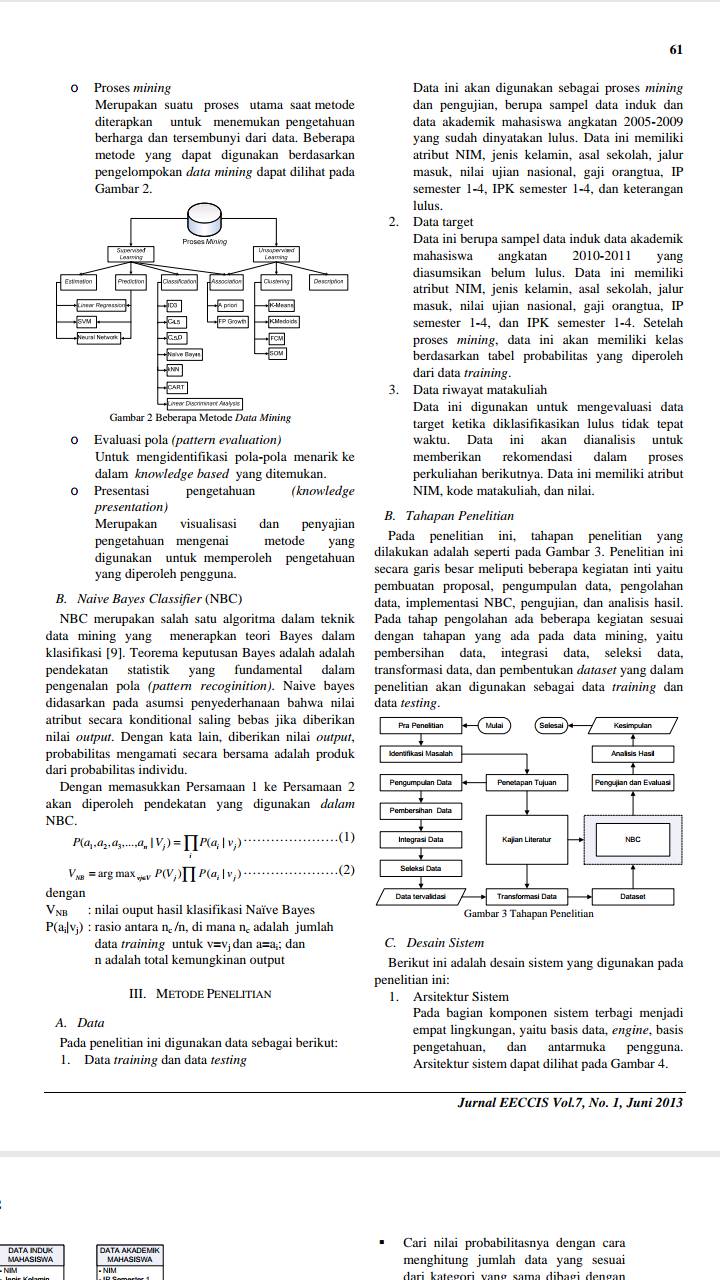
* Pemilihan tugas data mining; pemilihan goal dari proses KDD misalnya klasifikasi, regresi, clustering, dll.
* Pemilihan algoritma data mining untuk pencarian (searching)
* Proses Data mining yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

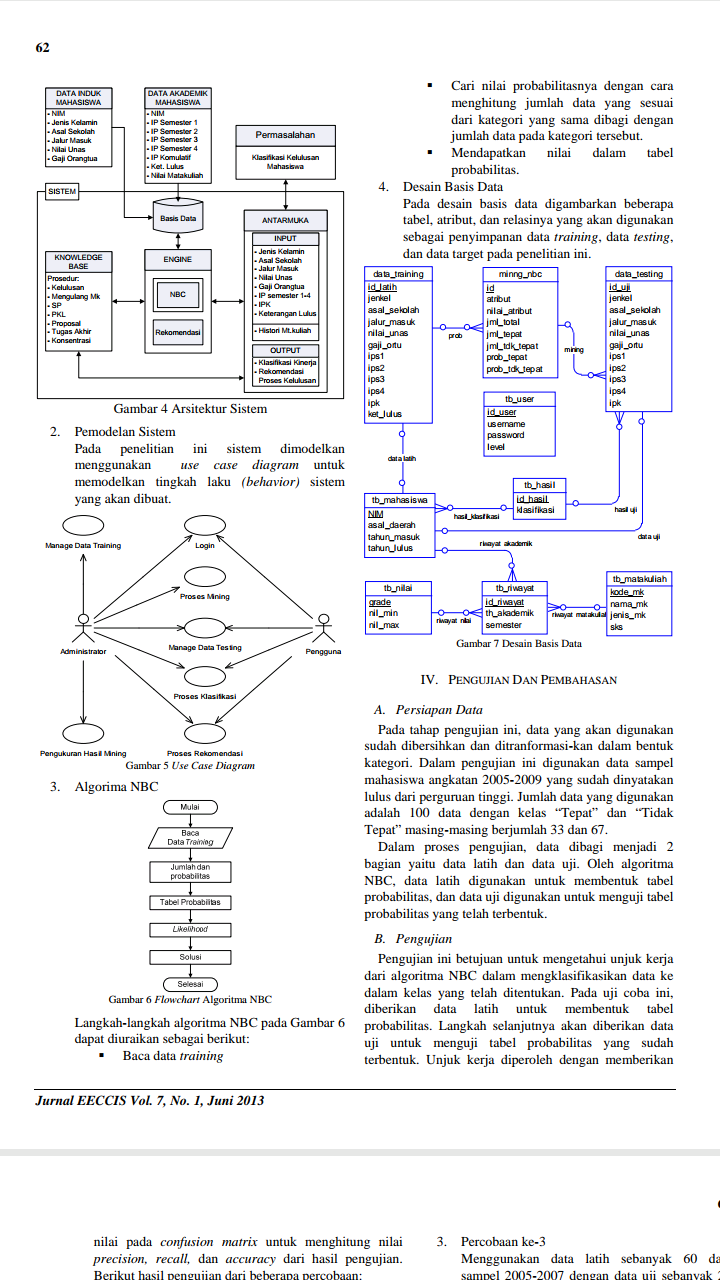
**5. Interpretation/ Evaluation**

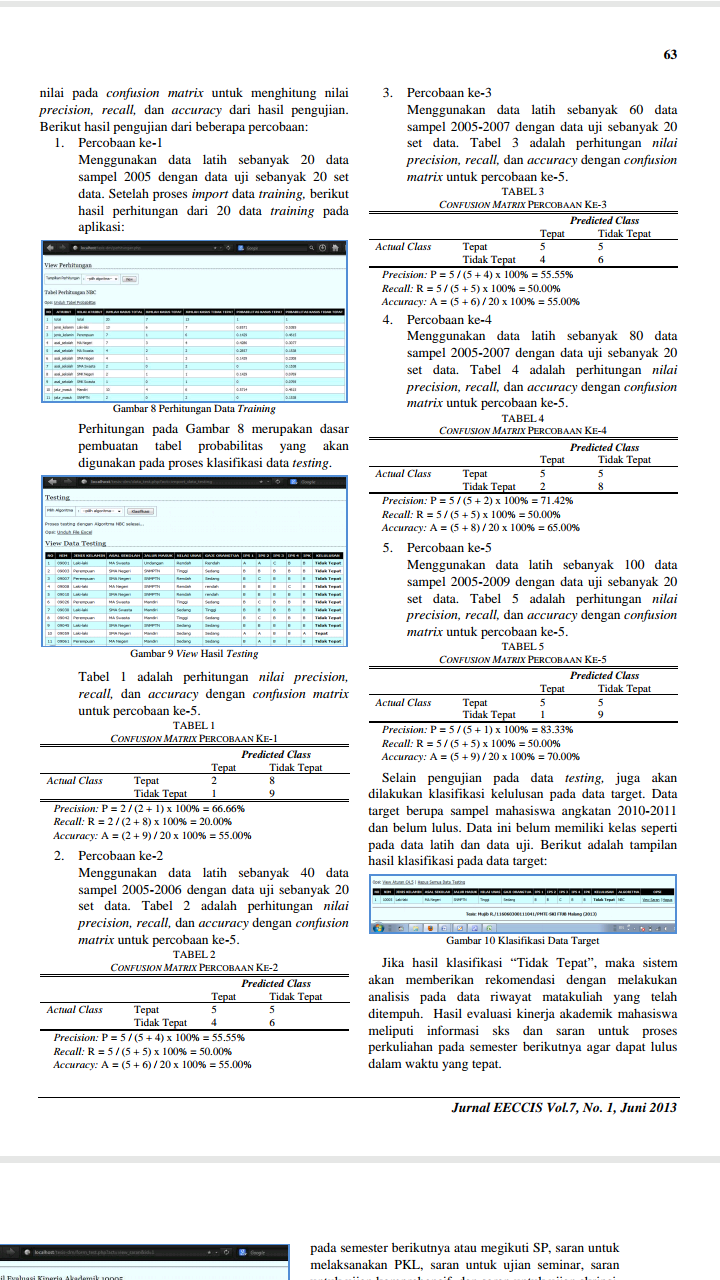
* Penerjemahan pola-pola yang dihasilkan dari data mining.
* Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah mimengerti oleh pihak yang berkepentingan.
* Tahap ini merupakan bagian dari proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

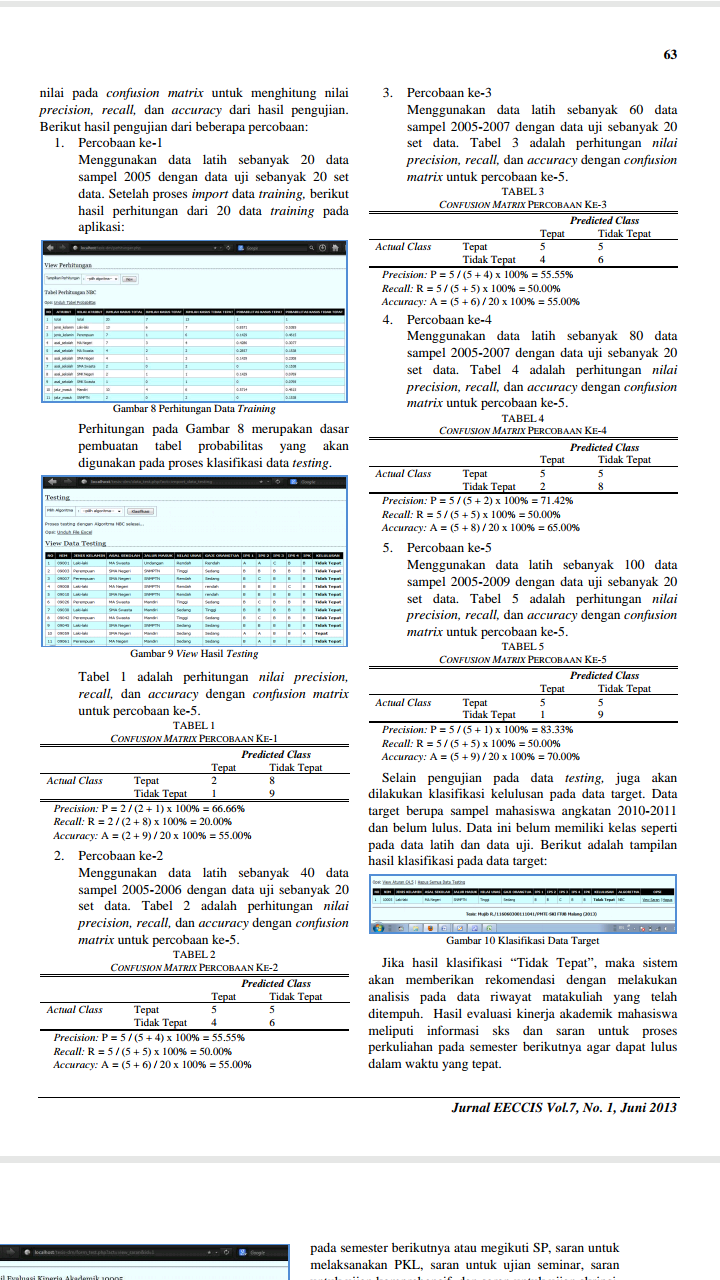












Analisis:

Dalam penerapan data mining untuk evaluasi kinerja akademik mahasiswa menggunakan algoritma naïve bayes classifer dapat disimpulkan bahwa penelitian tersebut menghasilkan suatu metode yang mana metode dapat membantu peneliti megevaluasi dan memberi solusi bagaimana cara nya memandu mahasiswa lulus dengan waktu yang tepat dan nilai yang optimal berdasarkan histori nilai yang ditempu dari metode tersebut akan membentuk sutu probabilitas sebagai dasar proses klasifikasi kelulusan mahasiswa dengan output berupa klasifikasi kinerja akademikdimana hasil pengujian menunjukan bahwa factor yang paling berpengaruh dalam klasifikasi kelulusan mahasiswa adalah indeks prestasi komulatif(ipk), indeks prestas(ip) dan jenis kelamin dari faktor tersebut dapat digunakan sebagai bahan evaluasi yang mana hasil dari pengujian evaluasi dapat membantu dan memberi solusi apa saja yang dapat membuat mahasiswa lulus dengan tepat waktu. Dari penentuan data training dapat mempengaruhi hasil pengujian karena pola data training tersebut akan dijadikan sebagi rule untuk menentukan kelas pada testing sehingga besar atau kecilnya persentase tingkat precision, recall, dan accuracy dipengaruhi juga oleh penentuan data training, hasil dari mining nbc juga dapat digunakan untuk mengklasifikasi kinerja akademik mahasiswa tahun ke 2 yang dalam penelitian ini dijadikan data target.