

## Penyusunan Prototype Sistem Informasi Pemetaan SDM-IG Berbasis Online

Desi Indriani Lestari<sup>1</sup>, Puspa Eosina<sup>2</sup>, Freza Riana<sup>3</sup>, Budi Susetyo<sup>4\*</sup>

<sup>1</sup>Jurusan Teknik Informatika, Fakultas Teknik, Universitas Ibn Khaldun Bogor

<sup>2</sup>Jurusan Teknik Informatika, Fakultas Teknik, Universitas Ibn Khaldun Bogor

<sup>3</sup>Jurusan Teknik Informatika, Fakultas Teknik, Universitas Ibn Khaldun Bogor

<sup>4</sup>Jurusan Teknik Informatika, Fakultas Teknik, Universitas Ibn Khaldun Bogor

Jl. Baru, Kedung Badak, Tanah Sareal, Kota Bogor - Jawa Barat

E-mail: Desiindriani08b@gmail.com 1), puspa.eosina@ft.uika-bogor.ac.id@ 2), freza@uika-bogor.ac.id 3), budi.susetyo@ft.uika-bogor.ac.id 4\*)

### ABSTRAKS

Perkembangan informasi geospasial di Indonesia semakin pesat, seiring dengan kebutuhan bisnis pengembangan aplikasi dan penyediaan jasa layanan informasi spasial tersebut. UU No. 4 tahun 2011 tentang informasi geospasial telah mengamanatkan penyediaan data dan akses terhadap informasi geospasial yang sangat diperlukan bagi proses pembangunan serta dasar perencanaan penataan ruang, kebencanaan, pengelolaan sumber daya alam, dan sumber daya lainnya, sehingga dapat dipertanggung jawabkan dan bermanfaat bagi masyarakat Indonesia. Secara umum dapat dikatakan bahwa kebutuhan SDM IG yang memiliki kompetensi dibidang IG ini semakin meningkat. Oleh karena itu dibutuhkan suatu sistem pemetaan kompetensi SDM IG berbasis web yang dapat mengarahkan dan memberikan rekomendasi peningkatan kompetensi yang sesuai dengan keterampilan yang dimilikinya. Data awal SDM IG ini merupakan data primer yang diperoleh dari hasil kuesioner online yang diisi langsung oleh SDM IG melalui website tersedia. Dalam melaksanakan penelitian ini telah digali segala sesuatu yang berkaitan dengan kriteria kompetensi SDM IG sesuai dengan SKKNI (Standar Kualifikasi Kerja Nasional Indonesia) dan kerangka kerja nasional Indonesia, yang merupakan skema okupasi bidang IG yang telah diterapkan oleh BIG dan BSN. Disamping data primer yang dijangkau melalui kuesioner online, juga diperoleh data sekunder dari pengisian kuesioner di basis penghasil SDM IG pada Lembaga Pendidikan tertentu di Indonesia. Tujuan dari penelitian ini adalah menyusun prototype sistem yang dapat memetakan taraf kompetensi SDM IG berdasarkan ketentuan dalam SKKNI dan KKNi bidang informasi geospasial. Kuesioner online ini digunakan untuk mengevaluasi SDM IG yang telah berlangsung setiap SDM mengisi kuesioner berdasarkan kompetensi yang ada. Data kuesioner yang telah diisi SDM dapat diakses melalui sistem informasi pemetaan SDM-IG berbasis online. Dengan adanya sistem diharapkan pengisian kuesioner dapat dilakukan dengan lebih cepat dan akurat. Sistem ini dibangun berbasis web. Metode pengembangan yang digunakan adalah waterfall, yaitu dimulai dengan menganalisa kebutuhan sistem, desain sistem, tahap penulisan kode program, pengujian sistem dan penerapan serta pemeliharaan sistem, lalu dalam tahap perancangan basis data. Dengan entitas dan attribute yang telah ditentukan, maka informasi yang dikelola ini dapat dimanfaatkan sebagai bahan perencanaan, pelaksanaan, evaluasi dan pemetaan kompetensi berdasarkan indeks kompetensi SDM yang dihasilkan dari pengolahan data mining.

**Kata Kunci:** teknologi informasi, industri geospasial, kompetensi, sdm-IG, kuesioner online

### 1. PENDAHULUAN

#### 1.1 Latar Belakang

Perkembangan teknologi industri Geospasial (IG) saat ini semakin hari semakin berkembang, seiring dengan perkembangan tersebut, berbagai perusahaan dan lembaga negeri maupun swasta mulai terlibat dalam bisnis pengembangan teknologi dan penyedia layanan. Dengan ini diperlukan pengetahuan dan pemahaman tentang kondisi industri geospasial. Pemerintah membuat UU No. 4 tahun 2011 tentang informasi geospasial telah disahkan pada tanggal 21 April

2011, yang mana undang-undang ini menjamin ketersediaan dan akses terhadap informasi geospasial yang sangat diperlukan untuk berbagai proses pembangunan dan dasar perencanaan penata ruang bencana, pengelolaan sumber daya alam, dan sumber daya lainnya, sehingga dapat dipertanggung jawabkan dan bermanfaat bagi masyarakat Indonesia (Nurhayati, Susetyo and Eosina, 2017).

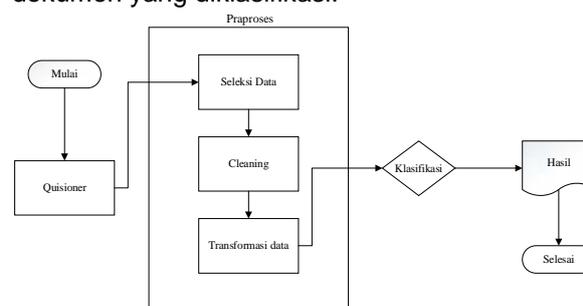
Kompetensi adalah karakteristik bahwa individu memiliki kemampuan dan digunakan dengan cara yang konsisten untuk mencapai kinerja yang diinginkan. Karakteristik ini meliputi pengetahuan, keterampilan, aspek citra diri, motif sosial, sifat, pola pikir dan cara berpikir, perasaan, dan pelaksanaan (Ediyanto, hal 133-136,2013).Perhitungan yang digunakan untuk mengelompokan penentuan kompetensi SDMIGdalam menghitung tiap hasil survei perusahaan adalah menggunakan Algoritma Klasifikasi Naive Bayes, algoritma Klasifikasi naive bayes memiliki tingkat ketelitian yang cukup tinggi dalam mengelompokan data. Sebuah Sistem berbasis web yang akan dibangun untuk memberikan wadah kepada para perusahaan untuk berkontribusi untuk kemajuan SDM IG di indonesia, yang dimana dari hasil penelitian dapat menentukan kompetensi indeks Sumber daya manusia (SDM)Adapun permasalahan yang dapat diangkat berdasarkan uraian pada latar belakang yaitu Bagaimana menganalisis potensi yang di butuhkan dalam proses penentuan kompetensi indeks SDM IG, Bagaimana menentukan metode yang dapat digunakan untuk menentukan tingkatan kompetensi, Bagaimana merancang dan mengimplementasikan sistem monitoring kompetensi indeks SDM IG dari hasil rancangan system, Adapun tujuan dari penelitian ini adalahMenganalisis potensi-potensi yang ada dan menentukan sebagai faktor yang berpengaruh dalam proses penentuan kompetensi indeks SDM IG,Menerapkan metode klasifikasi naive bayes dalam menentukan kompetensi tingkatan indeks SDM IG, Merancang dan membangun sistem berbasis web untuk menganalisis penentuan kompetensi indeks SDM IG.

### 1.2 Klasifikasi Naive Bayes

Klasifikasi adalah suatu fungsionalitas data mining yang akan menghasilkan model untuk memprediksi kelas atau kategori dari objek-objek di dalam basis data. Klasifikasi merupakan proses yang terdiri dari dua tahap, yaitu tahap pembelajaran dan tahap pengklasifikasian Pada tahap pembelajaran, sebuah algoritma klasifikasi akan membangun sebuah model klasifikasi dengan cara mengalisis training data. Tahap pembelajaran dapat juga dipandang sebagai tahap pembentukan fungsi atau pemetaan  $Y=F(X)$  dimana Y adalah kelas hasil prediksi dan X adalah tuple yang ingin diprediksi kelasnya. Selanjutnya pada tahap pengklasifikasian, model yang telah dihasilkan akan digunakan untuk melakukan klasifikasi.

*Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang

dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai teorema Bayes dengan asumsi indenpendensi (ketidaktergantungan) yang kuat (naif). Dengan kata lain, dalam *Naive Bayes* model yang digunakan adalah model dengan fitur indenpenden (Prasetyo, 2012). Algoritma *Naive Bayes* sendiri memiliki dua tahap dalam prosesklasifikasi, yaitu tahap pembelajaran dan tahap pengujian klasifikasi. Pada tahap pembelajaran dilakukan proses analisis terhadap sampel dokumen berupa pemilihan *vocabulary*, yaitu kata yang mungkin muncul dalam koleksi sampel dokumen yang sedapat mungkin dapat menjadi representasi dokumen. Selanjutnyaadalahpenentuanprobabilitaspriorbag itiapkategori berdasarkan sampel dokumen. Pada tahap klasifikasi ditentukan nilai kategori dari dokumen berdasarkan term yang muncul dalam dokumen yang diklasifikasi.



Gambar 1. Metode Penelitian

### 1.3 Studi Pustaka

| Judul penelitian   | Deskripsi  |
|--|--|
| Analisis Daya saing Industri Informasi Geospasial di Indonesia, Immas Nurhayati, Budi Susetyo, Puspa Eosina, Indu Pumahayu | Penelitian ini bertujuan untuk menganalisis penilaian daya saing industri IG untuk menghadapi MEA atau era pasar bebas yang akan datang, dan persaingan global dan pasar internasional |
| Implementasi Algoritma K-means Clustering untuk menentukan strategi marketing President                                    | Penelitian ini bertujuan untuk melihat strategi khusus oleh bagian marketing dalam melakukan pemasaran untuk mencari calon   |

---

|  |   |
|--|---|
| University, Johan Oscar Ong1.  | mahasiswa agar promosi yang dilakukan lebih efektif dan efisien. Untuk dapat melakukan promosi yang lebih efektif dan efisien, maka dalam penelitian ini dilakukan dengan cara mengolah data-data yang telah didapatkan dari mahasiswa yang telah lulus seperti nama mahasiswa, kota asal, jurusan yang diambil dan yang terakhir adalah nilai IPK. |
|  | Penelitian ini bertujuan untuk membagi tingkat kecanduan game online terhadap korelasinya terhadap prestasi akademik  |
| Pembagian Tingkat Kecanduan Game online menggunakan K-means clustering serta korelasinya terhadap prestasi akademik, Yudi Prasetyo |   |

---

## 2. PEMBAHASAN

Dalam penelitian ini, *prototipe* sistem mengklasifikasikan Kuesioner untuk pemetaan kompetensi SDM informasi geospasial dari data yang didapatkan dari SKKNI dan Kuesioner menggunakan algoritma naive bayes data ini berjumlah 55 dan dilakukan 10 kali percobaan untuk membandingkan hasil akurasi data SDM IG Sebelum data diolah menggunakan sistem yang akan dirancang, dilakukan metode pemrosesan data awal terlebih dahulu sesuai dengan proses KDD. Terdapat 4 langkah dalam melakukan pemrosesan data awal, yaitu pembersihan data (data cleaning), seleksi data, transformasi data, dan penambangan data.

### 2.1 Pembersihan Data

Pada tahap ini akan dilakukan pembersihan data yang tidak lengkap, kosong (null), *noise*, dan data yang tidak konsisten. Dalam langkah ini, data yang bernilai kosong (null), akan dibersihkan dengan cara dihapus secara manual, dan akan dilakukan penghapusan atribut atau mengganti data tersebut. Pembersihan data dilakukan setelah proses integrasi dan seleksi data, dikarenakan proses tersebut akan memudahkan proses pencarian data, sementara proses seleksi data akan mengurangi jumlah data yang akan dibersihkan.

### 2.2 Seleksi Data

Pada tahap ini, akan dilakukan penyelesaian data untuk mengurangi data yang tidak relevan, dan berlebihan (*redundant*). Menurut Tan, (Steinbach, dan Kumar 2006), atribut yang tidak relevan adalah atribut yang berisi informasi yang tidak berguna untuk melakukan penambangan data, sedangkan atribut yang berlebihan (*redundant*) adalah atribut yang menduplikasikan banyak atau semua informasi yang terdapat di dalam satu atau lebih pada atribut lainnya.

### 2.3 Transformasi Data

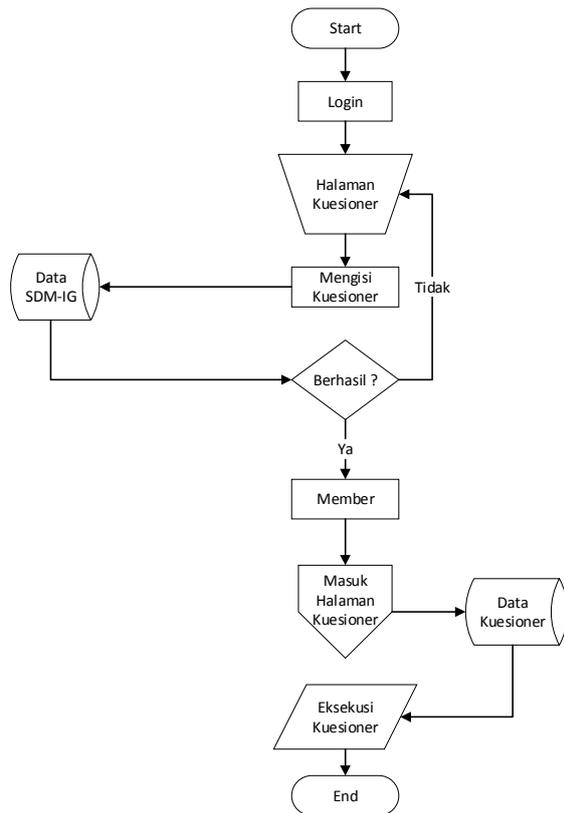
Pada tahap ini akan dilakukan transformasi data. Data Kuesioner diklasifikasikan menjadi beberapa jenis sesuai dengan jurnal ilmiah *Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure (JNC)*.

### 2.4 Penambangan Data

Data yang telah diproses pada tahap sebelumnya kemudian akan diuji menggunakan algoritma Naive Bayes. Langkah awal adalah proses input dimana sistem akan membaca data teks. Setelah data dibaca, data berupa string akan tercampur di dalam satu baris data, sehingga data perlu dipisah untuk kemudian diubah ke tipe data double sehingga data dapat diproses ke dalam perhitungan klasifikasi. Untuk melakukan pemrosesan data tersebut maka sistem perlu menulis ulang data teks kuesioner sehingga data teks kuesioner dapat diproses oleh sistem dan menghasilkan output berupa hasil akurasi dan prediksi jenis kuesioner.

### 2.5 Flowchart Kuesioner Online

Dalam perancangan sistem kuisisioner yang dibuat dapat digambarkan dengan diagram flowchart sebagai berikut:



Gambar 2. Flowchart Kuesioner Online

Dalam diagram alur flowchart kuisionerdapat dijelaskan user yang dapat masuk ke halaman kuisioner setelah mengisi form login dengan benar yaitu dengan memasukkan biodata. Setelah user dapat masuk ke halaman kuisioner user dapat mengisi form data kuisioner dan mengeksekusinya, setelah itu data yang telah dieksekusi langsung masuk ke dalam database kuisioner.

## 2.6 Hasil Perhitungan Naive Bayes

|                            |          |                |
|----------------------------|----------|----------------|
| P(Kompeten/Tidak Kompeten) | 56%      | 44%            |
|                            | 100%     |                |
|                            | Kompeten | Tidak Kompeten |
| P(Gender)                  | Kompeten | Tidak Kompeten |
| Pria                       | 52%      | 46%            |
| Wanita                     | 48%      | 54%            |
|                            | 100%     | 100%           |
|                            | Kompeten | Tidak Kompeten |
| P(Inisialisasi)            | Kompeten | Tidak Kompeten |
| Yes                        | 94%      | 0%             |
| No                         | 6%       | 100%           |
|                            | 100%     | 100%           |

Akurasi pengujian dinyatakan dalam persentase. Pada proses training (pelatihan) data digunakan untuk melatih sistem agar dapat mengenali dan menemukan bobot yang sesuai. Secara teori, dalam penelitian ini terdapat faktor yang dapat mempengaruhi tinggi rendahnya akurasi pengujian, faktor tersebut adalah banyaknya jumlah data latih yang digunakan dalam melatih sistem untuk mengenali data yang digunakan.

Jumlah data keseluruhan adalah 55, yang kemudian dibagi menjadi 2 jenis data, yaitu data latih dan data uji. Untuk mengetahui sejauh mana pengaruh jumlah data latih, dilakukan 10 kali proses pembelajaran dengan jumlah data latih yang berbeda,

| No | Banyak Data Total             | Rata-rata Akurasi (%) |
|----|-------------------------------|-----------------------|
| 1  | 40 data latih dan 15 data uji | 100                   |
| 2  | 45 data latih dan 10 data uji | 80                    |
| 3  | 30 data latih dan 25 data uji | 100                   |
| 4  | 37 data latih dan 18 data uji | 94                    |
| 5  | 28 data latih dan 27 data uji | 96                    |
| 6  | 46 data latih dan 9 data uji  | 100                   |
| 7  | 43 data latih dan 12 data uji | 92                    |
| 8  | 32 data latih dan 23 data uji | 100                   |
| 9  | 50 data latih dan 5 data uji  | 80                    |
| 10 | 35 data latih dan 20 data uji | 100                   |

Dari hasil diatas, terlihat bahwa semakin banyak jumlah data latih, maka akurasi pengujian akan semakin baik, sebaliknya, jika semakin sedikit jumlah data latih maka akurasi pengujian akan semakin buruk. Hal ini terjadi karena ketika data yang digunakan untuk training semakin banyak, maka sistem akan semakin memiliki banyak referensi dalam melakukan klasifikasi.

## 3. KESIMPULAN

Algoritma yang digunakan dalam sistem dapat mengklasifikasi kompeten dan tidak kompeten, dengan akurasi mencapai 100%. Parameter yang digunakan hanya pada jumlah data latih. Semakin banyak data latih yang digunakan, maka sistem yang digunakan akan semakin pintar dalam menentukan prediksi dari data uji yang digunakan.

### Ucapan Terima Kasih

Penulis mengucapkan terima kasih Tim Penelitian Strategi Nasional Universitas Ibn Khaldun Bogor yang telah memfasilitasi penelitian penunjang untuk Skema PSNI Ristek-Dikti dan Badan Informasi Geospasial (BIG) yang telah mendukung penelitian ini hingga selesai.

### PUSTAKA

Susetyo, B.,I. Nurhayati,P. Eosina. "Analisis Daya Saing Industri Informasi Geospasial di Indonesia," *Seminar nasional XII, Rekayasa Teknologi Industri dan Informasi*, 2017.

Ediyanto, "Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis,"

*Buletin Ilmiah Mat. Stat dan Terapannya (Bimaster)*, vol. 02, no. 2, pp. hal 133-136, (2013).

Jiawei Han, Micheline Kamber *Data Mining: concept and techniques*. (2006).

Natalius, S. 2010. *Metoda Naïve Bayes Classifier dan Penggunaanya pada Klasifikasi Dokumen*

Hamzah, A. 2012. *Klasifikasi Teks dengan Naïve Bayes Classifier (NBC) Untuk Pengelompokan Teks Berita dan AbstractAkademis*. ISSN: 1979-911X vol. 3., 2012