



BAB 1 PENDAHULUAN

Seiring bertambahnya waktu, tidak dapat dipungkiri bahwa peran teknologi informasi menjadi hal yang tidak dapat dipisahkan dalam kehidupan manusia terutama di era digital seperti sekarang. Jenis kebutuhan manusia yang beragam dan terus bertambah juga turut menjadi alasan mengapa komputasi digital menjadi hal yang terus berkembang setiap saat. Teknologi pengolahan data juga turut dituntut agar dapat selalu beradaptasi dalam mengimbangi kebutuhan manusia setiap harinya. Hal ini bertujuan agar bisa menghasilkan sebuah sistem yang dapat memproses data dalam jumlah yang banyak dalam waktu singkat. Dalam kehidupan manusia setiap hal atau sebuah kejadian yang terjadi tiap waktu mampu diterjemahkan sebagai sebuah fakta. Kemudian fakta tersebut dapat diolah kembali menjadi sebuah data yang selanjutnya dapat diproses menjadi sebuah informasi yang memiliki pola atau bentuk. Pola dan bentuk yang terdapat dalam informasi tersebut kemudian dapat diolah dalam sebuah sistem. Pengolahan informasi dalam jumlah yang besar tentunya membutuhkan sebuah sistem yang mampu melakukan proses komputasi dan pengolahan data yang akurat, cepat, serta efisien.

Berbagai macam jenis sistem yang telah berkembang dan hadir di kehidupan manusia sudah pasti bertujuan untuk mempercepat kinerja pengolahan informasi agar dapat menyelesaikan bermacam jenis masalah dalam kehidupan sehari-hari. Ketersediaan sistem pengolahan data yang baik tentunya akan mampu menyelesaikan berbagai masalah dan mengolah informasi dengan cepat serta efisien. Salah satu manfaat yang bisa kita peroleh dengan melakukan penerapan teknologi informasi dan sistem yang mumpuni adalah kemampuan untuk melakukan analisis dan prediksi data dalam jumlah yang sangat besar. Sejumlah data dengan volume yang besar dapat di analisis dan diterjemahkan menjadi



sebuah pola agar proses pengukuran dan prediksi nilai sesuai dengan kebutuhan di masa depan. Berdasarkan alasan tersebut, penulis kemudian memutuskan untuk melakukan penelitian untuk mengukur tingkat akurasi algoritma prediksi dengan sampel data Indeks Pembangunan Manusia (IPM) per provinsi di Indonesia.

Indeks Pembangunan Manusia (IPM) atau *Human Development Index (HDI)* adalah sebuah pengukuran dan perbandingan dari angka harapan hidup, angka melek huruf, tingkat pendidikan, dan standar kehidupan yang berlaku untuk semua negara di seluruh dunia. Nilai dari IPM ini kemudian digunakan untuk menentukan klasifikasi apakah sebuah negara bisa dikategorikan sebagai negara maju, negara berkembang, atau negara terbelakang. IPM juga bisa digunakan untuk mengukur pengaruh dari kebijakan ekonomi terhadap kualitas hidup manusia di sebuah negara. Menurut *Wikipedia*, penghitungan indeks ini dikembangkan oleh pemenang nobel India Amartya Sen dan seorang ekonom Pakistan Mahbub ul Haq, serta dibantu oleh Gustav Ranis dari Yale University dan Lord Meghnad Desai dari London School of Economics. Sejak saat itu indeks ini kemudian digunakan oleh PBB dalam program pembangunan mereka setiap tahun.

Seperti digambarkan oleh Amartya Sen. Indeks ini bisa disebut sebagai pengukuran vulgar karena berbagai batasan yang muncul. IPM lebih berfokus pada hal-hal yang lebih sensitif dan berguna dibandingkan hanya sekadar mengukur pendapat perkapita yang selama ini sering digunakan. IPM juga bisa menjadi sebuah jembatan untuk para peneliti yang lebih serius dalam meneliti hal-hal yang lebih terperinci dalam pembuatan laporan pembangunan manusia. Dalam IPM, pencapaian rata-rata sebuah negara bisa diukur dalam 3 dimensi dasar yaitu tingkat kehidupan yang sehat dan panjang umur yang diukur lewat angka harapan hidup saat lahir, tingkat pengetahuan yang diukur lewat tingkat baca tulis

orang dewasa, dan standar kelayakan hidup yang diukur lewat logaritma natural dari Produk Domestik Bruto (PDB) per kapita dalam partisipasi daya beli masyarakat.

Indonesia sebagai negara yang memiliki jumlah penduduk yang besar di dunia tentunya juga tidak terlepas dari pengukuran IPM dari PBB. Data IPM di Indonesia dikumpulkan dan dihitung melalui Badan Pusat Statistik (BPS) yaitu sebuah lembaga pemerintah non kementerian yang bertanggung jawab langsung kepada presiden untuk melakukan pengumpulan data tentang berbagai jenis bidang di Indonesia untuk kebutuhan pemerintah dan publik. Data mengenai IPM yang dikumpulkan tersebut meliputi data angka harapan hidup, angka melek huruf, lama sekolah serta pengeluaran perkapita. Data-data tersebut dikumpulkan berdasarkan setiap provinsi di Indonesia setiap tahun yang kemudian digunakan untuk menghitung IPM.

Berbagai data yang kemudian telah dikumpulkan oleh BPS untuk IPM dapat diakses melalui situs resmi BPS yaitu <http://bps.go.id> atau melalui portal data publik di Indonesia yaitu <http://data.go.id>. Data indeks IPM yang dikumpulkan oleh BPS memiliki tingkat akurasi yang tinggi karena dikumpulkan berdasarkan data tiap provinsi yang dilakukan dengan melakukan observasi langsung ke lapangan lewat tenaga ahli statistik dalam proses sehingga data tersebut sangat layak untuk digunakan dalam objek penelitian.

Lewat metode *data mining*, data IPM di Indonesia dapat diprediksi dengan menerapkan algoritma prediksi *Neural Network* dan *Generalized Linear Model*. Dengan menerapkan algoritma ini, kita dapat melakukan prediksi nilai IPM di masa depan secara sistematis dan cepat tanpa harus mengumpulkan variabel terkait seperti angka harapan hidup, angka melek huruf, lama sekolah atau pengeluaran perkapita yang membutuhkan waktu yang cukup lama. Dalam penelitian ini penulis juga akan melakukan pengukuran kelayakan kedua algoritma tersebut agar mampu memilih algoritma prediksi yang paling tepat untuk



mendapatkan data IPM di masa depan berdasarkan data pembanding dan sampel data yang telah disiapkan dari 2009 sampai dengan 2012.

Proses pengolahan data IPM secara manual tentu saja tidak menutup adanya berbagai masalah yang timbul. Penggunaan metode konvensional untuk pengumpulan data IPM secara manual yang memakan waktu sehingga data prediksi IPM tidak dapat diperoleh dengan cepat karena harus menunggu proses survey dan pendataan secara berkala yang kurang efektif. Pengumpulan data yang memakan waktu lama ini juga menyebabkan belum ditentukannya data acuan (*training*) yang tepat pada periode tertentu untuk pemodelan prediksi menggunakan metode *data mining*. Tanpa data acuan (*training*) tentunya akurasi algoritma prediksi *Neural Network* dan *Generalized Linear Model* yang belum terukur masih belum diketahui kelayakannya untuk memprediksi data IPM.

Selain itu, pengujian kelayakan algoritma tersebut tentu membutuhkan pembuktian jumlah variabel yang paling tepat untuk proses pengetesan. Karena perbandingan jumlah variabel untuk prediksi yang belum dilakukan sehingga tidak bisa mengetahui jumlah variabel yang paling tepat untuk menentukan akurasi algoritma prediksi *Neural Network* dan *Generalized Linear Model*. Hal ini juga berpengaruh karena laporan penghitungan akurasi algoritma prediksi berdasarkan variabel dan periode yang belum tercatat sehingga dokumentasi pengujian belum bisa dilakukan.

Berdasarkan hal-hal diatas, penulis kemudian memutuskan untuk melakukan penelitian ini yang bertujuan untuk melakukan pengujian dan pengukuran akurasi algoritma *Neural Network* dan *Generalized Linear Model* dengan variabel dan data periode masa lalu agar dapat melakukan prediksi nilai IPM secara akurat, efisien, dan mendekati data asli di masa depan.





B. Ruang Lingkup Penelitian

1. Identifikasi Masalah

Dari latar belakang yang dijelaskan, penulis kemudian mengidentifikasi beberapa masalah yang muncul seperti:

- a. Belum ada perhitungan menggunakan algoritma prediksi dan pengumpulan data IPM masih secara manual yang memakan waktu sehingga data prediksi IPM tidak dapat diperoleh dengan cepat karena harus menunggu proses survey dan pendataan secara berkala yang kurang efektif.
- b. Belum ditentukannya data acuan (*training*) yang tepat pada periode tertentu untuk pemodelan prediksi menggunakan metode *data mining*.
- c. Perbandingan jumlah variabel untuk prediksi yang belum dilakukan sehingga tidak bisa mengetahui jumlah variabel yang paling tepat untuk menentukan akurasi algoritma prediksi *Neural Network* dan *Generalized Linear Model*.
- d. Laporan penghitungan akurasi algoritma prediksi berdasarkan variabel dan periode yang belum tercatat sehingga dokumentasi pengujian belum bisa dilakukan.

2. Batasan Masalah

Maka berdasarkan identifikasi masalah yang telah disebutkan diatas penulis kemudian menentukan batasan-batasan dari masalah yang muncul yaitu:

- a. Belum ada perhitungan menggunakan algoritma prediksi dan pengumpulan data IPM masih secara manual yang memakan waktu sehingga data prediksi IPM tidak dapat diperoleh dengan cepat karena harus menunggu proses survey dan pendataan secara berkala yang kurang efektif.



- b. Belum ditentukannya data acuan (*training*) yang tepat pada pada periode tertentu untuk pemodelan prediksi menggunakan metode *data mining*.
- c. Perbandingan jumlah variabel untuk prediksi yang belum dilakukan sehingga tidak bisa mengetahui jumlah variabel yang paling tepat untuk menentukan akurasi algoritma prediksi *Neural Network* dan *Generalized Linear Model*.
- d. Laporan penghitungan akurasi algoritma prediksi berdasarkan variabel dan periode yang belum tercatat sehingga dokumentasi pengujian belum bisa dilakukan.

C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian Gie

C. Tujuan Penelitian

Berdasarkan batasan-batasan masalah yang telah ditentukan diatas, penulis memutuskan melakukan penelitian ini dengan tujuan untuk:

1. Menghemat waktu pengumpulan data IPM dengan metode *data mining* agar data IPM bisa diperoleh tanpa pengumpulan data secara manual lewat survey dan pendataan berkala yang kurang efektif.
2. Menentukan data acuan (*training*) yang tepat pada periode tertentu untuk menentukan pemodelan prediksi menggunakan metode *data mining*.
3. Membandingkan jumlah variabel prediksi agar jumlah variabel yang paling tepat bisa ditemukan untuk menentukan akurasi algoritma prediksi.
4. Membuat laporan penghitungan akurasi algoritma prediksi *Neuranl Network* dan *Generalized Linear Model* lewat rancangan *Graphical User Interface (GUI)* untuk dokumentasi pengujian yang bisa kembali sebagai acuan atau referensi penelitian di masa depan.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



D. Manfaat Penelitian

1. Bagi pembaca:

- a. Mendapatkan informasi mengenai tingkat akurasi algoritma *Neural Network* dan *Generalized Linear Model* untuk memprediksi nilai IPM dengan variabel dan periode data yang terbaik.
- b. Mampu memahami tingkat kesalahan (*error margin*) dari data IPM yang dihasilkan dari metode prediksi *data mining* dibandingkan dengan menggunakan metode konvensional.
- c. Menyediakan laporan perbandingan akurasi algoritma *Neural Network* dan *Generalized Linear Model* dalam bentuk GUI sebagai dokumentasi dan referensi penelitian lainnya.

2. Bagi Badan Pusat Statistik (BPS) dan Data.go.id:

- a. Menyediakan hasil penelitian pengukuran algoritma *Neural Network* dan *Generalized Linear Model* dengan data aktual yang terpercaya untuk membantu memprediksi nilai IPM di masa depan berdasarkan data aktual tahun sebelumnya.
- b. Menyediakan laporan perbandingan data IPM aktual dan prediksi untuk studi dan referensi penelitian di BPS atau Data.go.id.

3. Bagi penulis:

- a. Menerapkan berbagai ilmu dan pemahaman tentang *data mining*, *Neural Network* dan *Generalized Linear Model* dalam analisa, prediksi, dan pengukuran akurasi.
- b. Menggunakan metode pemodelan *data mining* untuk menyelesaikan masalah prediksi data dalam kehidupan nyata.

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.

- c. Meningkatkan pemahaman dan melakukan penerapan Ilmu atau konsep Sistem Informasi selama mengikuti perkuliahan di Institut Bisnis dan Informatika Kwik Kian Gie.

© Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

Institut Bisnis dan Informatika Kwik Kian Gie

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.

