

# BAB I

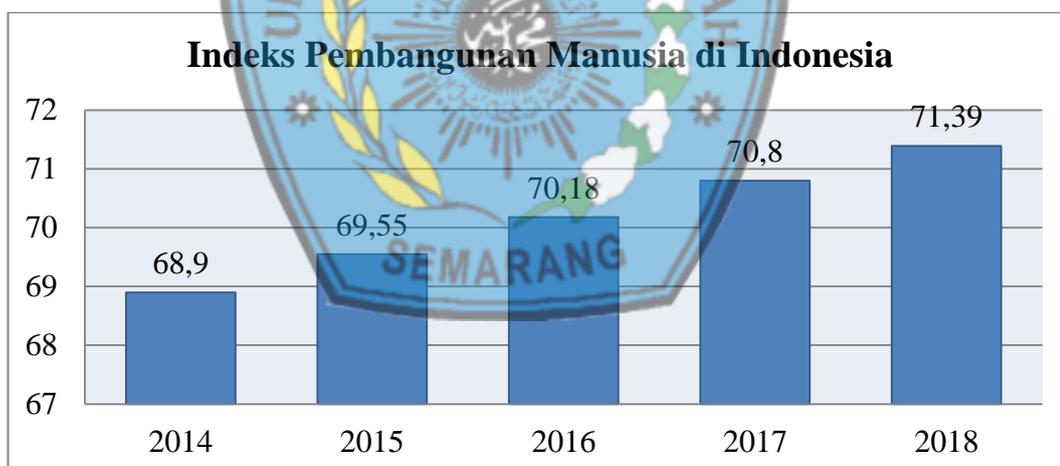
## PENDAHULUAN

### 1.1 Latar Belakang

Keberhasilan pembangunan nasional tidak hanya dilihat dari laju pertumbuhan ekonomi yang tinggi, tetapi yang paling penting adalah keberhasilan pembangunan manusia. Pembangunan manusia didefinisikan sebagai suatu proses untuk perluasan pilihan yang lebih banyak kepada penduduk melalui upaya-upaya pemberdayaan yang mengutamakan peningkatan kemampuan dasar manusia agar dapat sepenuhnya berpartisipasi disegala bidang pembangunan (Chalid, 2014).

Pembangunan manusia sebagai ukuran kinerja pembangunan secara keseluruhan dibentuk melalui pendekatan tiga dimensi dasar, yaitu umur panjang dan sehat, pengetahuan, dan penghidupan yang layak. Indikator-indikator yang mempresentasikan ketiga dimensi ini terangkum dalam satu nilai tunggal, yaitu angka Indeks Pembangunan Manusia (BPS, 2008). Indeks Pembangunan Manusia (IPM) atau disebut juga dengan *Human Development Index* (HDI) ditetapkan oleh Perserikatan Bangsa-Bangsa (PBB) sebagai ukuran standar pembangunan manusia. Secara global, IPM juga dapat digunakan sebagai penentu suatu negara termasuk dalam negara maju ataupun negara berkembang (Trianggara dkk, 2016). Untuk mengukur dimensi kesehatan, digunakan angka harapan hidup waktu lahir. Angka harapan hidup merupakan tolak ukur kualitas kesehatan di daerah tersebut, angka harapan hidup juga digunakan sebagai tolak ukur rata-rata masa hidup

daerah atau wilayah tersebut. Selanjutnya untuk mengukur dimensi pengetahuan digunakan gabungan indikator harapan lama sekolah dan rata-rata lama sekolah. Faktor pendidikan juga sangat penting dalam mendukung pembangunan manusia karena dari pendidikan manusia dapat mengembangkan dan berfikir maju. Sedangkan untuk mengukur dimensi hidup layak digunakan indikator kemampuan daya beli masyarakat terhadap sejumlah kebutuhan pokok yang dilihat dari rata-rata besarnya pengeluaran per kapita disesuaikan. Klasifikasi Indeks pembangunan manusia menurut Badan Pusat Statistika (2014) dikatakan rendah jika  $IPM < 60$ , sedang  $60 \leq IPM < 70$ , tinggi  $70 \leq IPM < 80$ , dan  $\geq 80$  sangat tinggi. IPM di Indonesia dalam lima tahun terakhir ini mengalami kenaikan, dapat dilihat pada gambar 1.1 berikut :



**Gambar 1.1** IPM di Indonesia

Berdasarkan gambar diatas mengenai Indeks Pembangunan Manusia di Indonesia dari tahun 2014 hingga tahun 2018 mengalami kenaikan yang cukup signifikan. Pada tahun 2014 nilai IPM tercatat 68.9, 69.55 untuk IPM tahun 2015, 70.18 untuk IPM tahun 2016, 70.8 untuk IPM tahun 2017 dan 71.39 untuk IPM

tahun 2018.

*Artificial Intelligence* (kecerdasan buatan) yang membuat mesin mampu menjawab pertanyaan dengan cerdas sekarang ini dalam bidang teknik berkembang sangat luas. *Machine Learning* memberi kita teknik terbaik untuk kecerdasan buatan seperti klasifikasi, regresi, pembelajaran yang diawasi, dan pembelajaran tanpa pengawasan, dan banyak lagi. Karena ini adalah masalah optimalisasi, maka banyak teknik optimalisasi juga digunakan untuk mengurangi dimensi dataset (kumpulan data) (Siburian, 2018).

Metode klasifikasi yang menggunakan teknik pohon keputusan yang populer umum digunakan yaitu *Classification and Regression Trees (CART)*, baik dalam penyelesaian kasus klasifikasi maupun regresi yang mana masuk dalam kelas *supervised learning*, dimana dalam penerapannya perlu adanya tahap *training* dan tahap *testing*. CART mampu menghasilkan klasifikasi yang akurat dan lebih cepat penghitungannya. Namun CART yang merupakan pohon keputusan tunggal cenderung memiliki model kurang stabil, karena perubahan sedikit saja pada data akan berpengaruh pada hasil prediksi dan cenderung *overfitting*. Oleh karena itu, untuk mengatasinya digunakan metode *ensemble* yaitu *Random Forest* yang mampu meningkatkan kestabilan pohon klasifikasi yang terbentuk dan akurasi prediksi yang dihasilkan (Seftiana, 2014). Metode *Random Forest* juga memiliki keunggulan untuk menghindari *overfitting* (Dhawangkhar, 2017). Metode *Random Forest* ini merupakan metode pengembangan dari metode CART yaitu dengan menerapkan metode *bootstrap aggregating bagging* dan *random feature selection* (Aliady, 2018).

Penelitian terkait Indeks Pembangunan Manusia (IPM) pernah dilakukan oleh Fauzi (2017) dengan judul “*Smooth Support Vector Machine (SSVM)* untuk Pengklasifikasian Indeks Pembangunan Manusia Kabupaten/Kota Se-Indonesia. Selain itu juga ada penelitian dengan judul “Pemodelan Indeks Pembangunan Manusia Menggunakan *Spatial Panel Fixed Effect*” dengan studi kasus IPM Provinsi Jawa Tengah 2008-2013 (Trianggara dkk, 2016).

Penelitian terkait *Random Forest* pernah dilakukan oleh (Dhawangkharu & Riksakomara, 2017) dengan judul “Intensitas Hujan Kota Surabaya dengan Menggunakan teknik *Random Forest* dan CART.” Kesimpulan yang didapat dari penelitian yang dilakukan bahwa Metode *Random Forest* memiliki performa sedikit lebih baik dibandingkan dengan CART sebesar 6%. Sedangkan penelitian yang dilakukan (Walinder, 2014) dengan judul “Evaluation of Logistic Regression and *Random Forest* Classification Based on Prediction Accuracy and Metadata Analysis”, menunjukkan bahwa metode *Random Forest* lebih baik hasil akurasi dalam pengklasifikasian data set. Penelitian lainnya mengenai *Random Forest* dilakukan (Yahya, 2018) dengan judul “Klasifikasi Ketepatan Lama Studi Mahasiswa Menggunakan Metode *Support Vector Machine* dan *Random Forest*” menunjukkan bahwa metode terbaik untuk menentukan ketepatan lama studi mahasiswa Universitas Islam Indonesia adalah *Random Forest*.

Dari penelitian diatas dapat disimpulkan bahwa *Random Forest* untuk mengatasi masalah data yang besar dan masih menghasilkan keakuratan yang baik. Jumlah kabupaten/kota di Indonesia terdata sangat banyak dengan jumlah sebesar 514 kabupaten/kota, dengan empat tingkatan IPM yaitu rendah, sedang,

tinggi, dan sangat tinggi. Untuk memprediksi IPM termasuk rendah, sedang, tinggi, dan sangat tinggi dengan melibatkan atribut pendukung yaitu angka harapan hidup, harapan lama sekolah, rata-rata lama sekolah, dan pengeluaran perkapita yang disesuaikan. Dengan jumlah kabupaten/kota se-Indonesia sebanyak 514, metode CART kurang baik untuk memprediksi serta menghasilkan keakuratan yang kurang baik, maka diperlukan metode khusus untuk menyelesaikan kasus yang mempunyai data berjumlah yang banyak, berdasarkan latar belakang yang dijelaskan diatas bahwa metode *Random Forest* sangat sesuai untuk kasus yang datanya berjumlah banyak dan belum adanya penelitian tentang klasifikasi IPM dengan metode *Random Forest* maka akan dilakukan penelitian mengenai klasifikasi IPM di Indonesia berdasarkan kabupaten/kota di Indonesia dengan metode *Random Forest*.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka permasalahan yang dirumuskan adalah sebagai berikut :

1. Bagaimana statistika deskriptif Indeks Pembangunan Manusia (IPM) kabupaten/kota di Indonesia tahun 2018 ?
2. Bagaimana prediksi dan akurasi hasil klasifikasi Indeks Pembangunan Manusia (IPM) kabupaten/kota di Indonesia tahun 2018 dengan metode *Random Forest* ?
3. Apakah *importance variable* yang sangat berpengaruh pada Indeks Pembangunan Manusia Kabupaten/Kota di Indonesia tahun 2018

dengan menggunakan metode *Random Forest* ?

### 1.3 Tujuan Penelitian

Tujuan dari penelitian ini berdasarkan masalah diatas adalah sebagai berikut :

1. Mendapatkan statistika deskriptif dari Indeks Pembangunan Manusia (IPM) kabupaten/kota di Indonesia tahun 2018.
2. Mengetahui prediksi dan akurasi hasil klasifikasi Indeks Pembangunan Manusia (IPM) kabupaten/kota di Indonesia tahun 2018 dengan metode *Random Forest*.
3. Mengetahui *importance variable* yang sangat berpengaruh pada Indeks Pembangunan Manusia Kabupaten/Kota di Indonesia tahun 2018 berdasarkan *Random Forest*.

### 1.4 Manfaat Penelitian

Penelitian ini memiliki beberapa manfaat diantaranya adalah sebagai berikut :

1. Manfaat Teoritis

Adanya penelitian ini diharapkan dapat memperluas wawasan dan menambah ilmu tentang teknik penyelesaian kasus dengan data yang berjumlah banyak dengan cara yang berbeda dari biasanya yaitu dengan menggunakan *Random Forest*.

## 2. Manfaat Praktis

- a. Bagi peneliti, peneliti mampu menerapkan metode sesuai dalam materi yang telah dipelajari serta peneliti mendapatkan pengetahuan dan wawasan mengenai klasifikasi menggunakan *Random Forest*.
- b. Bagi pihak terkait diantaranya pemerintah dan masyarakat yaitu dapat membantu mengklasifikasikan Indeks Pembangunan Manusia Kabupaten/Kota di Indonesia agar bisa melihat/memprediksi IPM dengan tingkatan rendah, sedang, tinggi, dan sangat tinggi. Dengan mengetahui kemungkinan IPM rendah, maka pemerintah dapat melakukan tindakan untuk mengatasinya.

### 1.5 Batasan Penelitian

Implementasi dan analisis dalam metode *Random Forest* ini menggunakan *Software R*. Data yang digunakan adalah Indeks Pembangunan Manusia tahun 2018. Pengukuran ketepatan klasifikasi diukur melalui *total accuracy rate* (1-APER) yang menunjukkan akurasi keseluruhan suatu klasifikasi.