

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kesehatan merupakan aspek esensial dan aspek yang sangat diperjuangkan dalam kehidupan manusia. Dalam memperjuangkan aspek kesehatan pada suatu negara, sangatlah lazim jika negara tersebut memberikan tanggungan dan jaminan bagi warga negaranya. Oleh karena itu, keharusan negara untuk menyediakan jaminan sosial dan kesehatan bagi warga negara tertuang pada UUD 1945 pasal 34. Melihat dari sejarah perkembangan jaminan sosial Indonesia yang awalnya hanya mencakup pegawai negeri sipil hingga mencakup seluruh warga negara, hal ini menandakan bahwa jaminan sosial adalah hal yang sangat penting untuk dimiliki oleh setiap warga negara.

Indonesia memberikan jaminan kesehatan bagi seluruh warga negaranya melalui program Jaminan Kesehatan Nasional (JKN) yang diselenggarakan oleh suatu badan bernama Badan Penyelenggara Jaminan Sosial (BPJS). Melalui data yang ditampilkan oleh situs resmi BPJS Kesehatan pada 31 Mei 2020, tercatat bahwa jumlah peserta program JKN adalah 220.687.267 orang atau sekitar 81,85% dari seluruh rakyat Indonesia. Walaupun hampir seluruh warga negara Indonesia sudah terdaftar pada program JKN, pada praktiknya BPJS Kesehatan mengalami keuangan yang defisit. Sesuai keterangan yang diberikan oleh Menteri Keuangan Sri Mulyani pada 9 Maret 2020 lalu, keuangan BPJS Kesehatan mencapai defisit

dengan nominal 13 triliun Rupiah walaupun sudah diberikan suntikan dana sebesar 15 triliun Rupiah.

Untuk menanggulangi permasalahan ini, BPJS Kesehatan merencanakan kenaikan iuran tiap kelas sebesar hampir dua kali lipat pada setiap kelasnya. Rencana ini tentu menimbulkan penolakan dari masyarakat. Beberapa pihak menilai bahwa langkah tersebut bukanlah langkah yang baik. Salah satu pihak yang berpendapat demikian adalah Komisi Pemberantasan Korupsi (KPK). KPK berpendapat bahwa defisit keuangan tersebut disebabkan oleh *fraud* (penyimpangan) dan inefisiensi sistem sehingga menaikkan iuran tidak akan menyembuhkan BPJS Kesehatan dari defisitnya.

Perkembangan teknologi pada era Industri 4.0 membuka peluang besar dalam akuisisi data dan pengolahannya. Teknik *data mining* merupakan suatu teknik yang semakin populer dan mulai berubah menjadi suatu hal esensial terutama pada sistem lembaga kesehatan. *Data mining* dapat diartikan sebagai proses untuk mencari pola dan tren yang sebelumnya tidak terlihat pada data yang berukuran sangat besar [1]. Walaupun bukan suatu hal yang baru, *data mining* masih sangat berguna bagi banyak institusi hingga saat ini untuk berbagai pekerjaan termasuk *fraud detection* dan digunakan sebagai *data-driven decision making*. Akhir – akhir ini, teknik *data mining* sering menggunakan algoritma *machine learning*, contohnya seperti algoritma *K-Means Clustering*.

Dengan dirilisnya data sampel dari pengguna BPJS Kesehatan periode 2015 – 2016, BPJS Kesehatan membuka peluang bagi para peneliti untuk melakukan

penelitian lebih lanjut. Dengan begitu, peneliti dapat mengetahui seluk beluk dari permasalahan yang muncul ke publik dengan cara mengolah data tersebut. Pada penelitian ini, penulis akan mengolah data tersebut baik secara deskriptif maupun prediktif. Prediksi yang akan dilakukan pada penelitian ini akan berkisar pada prediksi jumlah klaim yang akan dilakukan pada rentang bulan tertentu dengan cara membentuk model matematika dan *machine learning*.

Penelitian ini juga diharapkan dapat menjawab dan menjelaskan apa yang terjadi pada sistem BPJS Kesehatan dengan cara mengolah data sampel tersebut lalu mengelompokkan sampel ke dalam beberapa klaster sesuai dengan kemiripannya sehingga terlihat secara jelas perilaku dari setiap kelompok sampel. Setelah terpisah, klaster – klaster tersebut akan dianalisis sehingga dapat ditarik kesimpulan tentang klaster apa yang menyebabkan kerugian pada sistem BPJS Kesehatan Indonesia.

1.2. Maksud dan Tujuan

Skripsi ini dibuat dengan tujuan untuk menganalisis sampel – sampel pengguna BPJS Kesehatan Indonesia dan mengelompokkannya menjadi beberapa klaster sesuai kemiripan sampel sehingga dapat dihasilkan sebuah analisis yang bersifat deskriptif. Analisis deskriptif tersebut diharapkan dapat mencerminkan tentang apa yang terjadi pada sampel – sampel tersebut sehingga dapat membantu pihak BPJS Kesehatan Indonesia sebagai dasar dalam mengambil keputusan demi menurunkan defisit yang sedang dialami oleh BPJS Kesehatan.

Selain itu, skripsi ini bertujuan untuk melakukan analisis prediktif untuk memprediksi jumlah klaim yang akan terjadi. Hasil prediksi tersebut dibentuk menggunakan model matematika dan *machine learning*, sehingga diharapkan nantinya model tersebut dapat digunakan untuk memprediksi periode – periode selanjutnya dan pihak BPJS Kesehatan Indonesia dapat melakukan tindakan preventif terhadap *outcome* yang akan muncul.

1.3. Batasan Masalah

Skripsi ini memiliki batasan masalah sebagai berikut :

1. *Dataset* yang digunakan adalah data Kepesertaan pengguna layanan BPJS Kesehatan, data Kunjungan pada Fasilitas Kesehatan Tingkat Pertama Nonkapitasi (FKTP Nonkapitasi), dan data Kunjungan pada Fasilitas Kesehatan Rujukan Tingkat Lanjut (FKRTL) dengan rentang waktu dari 1 Januari 2015 hingga 31 Desember 2016.
2. Metode klusterisasi yang digunakan adalah klusterisasi dengan menggunakan algoritma *machine learning K-Means Clustering*.
3. Model *machine learning* yang digunakan untuk melakukan prediksi adalah *Random Forest Regressor*, *Linear Regression*, dan *ARIMA*. Selain itu, untuk melakukan prediksi digunakan juga model matematika untuk teori kredibilitas, yaitu model *non parametric Bühlmann-Straub credibility theory*.
4. Bahasa pemrograman yang digunakan adalah *Python 3* dengan *library machine learning Scikit-learn* dan *Statmodels*.

1.4. Metode Penelitian

Langkah – langkah yang dilakukan dalam penelitian ini adalah studi literatur, integrasi *dataset*, *data cleansing*, *Exploratory Data Analysis (EDA)*, *Recency-Frequency-Monetary (RFM) analysis*, klasterisasi sampel dengan algoritma *K-Means Clustering*, pembentukan model prediktif, dan evaluasi performa model. Pemaparan hasil klasterisasi dan prediksi kemudian divisualisasikan dengan grafik dan tabel.

1.5. Sistematika Penelitian

Skripsi ini ditulis dalam enam bab dengan sistematika sebagai berikut :

1. Bab I : Pendahuluan

Bab ini berisikan tentang latar belakang penelitian, tujuan penelitian, batasan masalah, metode penelitian dan sistematika penulisan.

2. Bab II : Landasan Teori

Bab ini memuat dasar teori dari penelitian ini. Bab ini berisikan teori mengenai *data science*, *descriptive analytics*, *predictive analytics*, *Python 3*, *machine learning*, algoritma *supervised & unsupervised*, algoritma *K-Means Clustering*, *non parametric Bühlmann-Straub credibility theory*, regresi dengan *Linear Regression*, *Random Forest Regressor*, *ARIMA*, penjelasan sistem BPJS Kesehatan, perubahan kebijakan pada BPJS Kesehatan, dan kebijakan BPJS Kesehatan pada tahun 2020.

3. **Bab III : Metode Penelitian**

Bab ini menjelaskan tentang langkah kerja beserta detail dari keseluruhan sistem yang digunakan untuk penelitian ini secara urut.

4. **Bab IV : *Exploratory Data Analysis***

Bab ini akan menggambarkan data yang digunakan untuk penelitian ini secara garis besar dan melihat trend dari data menggunakan grafik terhadap variabel – variabel tertentu.

5. **Bab V : Analisis dan Pembahasan**

Bab ini adalah bab yang akan menampilkan hasil prediksi menggunakan beberapa model yang telah dibuat dan mengevaluasi performa dari setiap modelnya. Selain itu, hasil prediksi akan ditampilkan dalam bentuk grafik sehingga dapat terlihat perbedaan antara hasil prediksi dan nilai sesungguhnya.

6. **Bab VI : Hasil dan Kesimpulan**

Bab ini memaparkan kesimpulan dari penelitian sekaligus saran untuk penelitian lanjut.