

BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Perangkat lunak menjadi salah satu teknologi terpenting di dunia yang tertanam dalam semua jenis sistem seperti transportasi, medis, telekomunikasi, militer, industri, dan hiburan. Seiring dengan semakin pentingnya perangkat lunak, komunitas perangkat lunak terus berupaya mengembangkan teknologi yang akan membuatnya lebih mudah, lebih cepat, dan lebih murah untuk membangun dan memelihara perangkat lunak yang berkualitas tinggi (Pressman & Maxim, 2015). Pada implementasinya, mengembangkan sebuah perangkat lunak yang berkualitas tinggi sulit dilakukan karena semakin kompleks fitur yang dimiliki suatu perangkat lunak, maka akan semakin banyak jumlah kode yang dihasilkan yang kemudian menurunkan kualitas perangkat lunak (Lacerda dkk., 2020).

Kualitas perangkat lunak merupakan kemampuan produk perangkat lunak untuk memenuhi kebutuhan yang dinyatakan secara langsung maupun tidak langsung berdasarkan kondisi tertentu (IEEE Computer Society, 2014). Terdapat delapan karakteristik dari kualitas perangkat lunak menurut ISO/IEC 25010:2011, yaitu kesesuaian fungsional (*functional suitability*), efisiensi performa (*performance efficiency*), kompatibilitas (*compatibility*), kegunaan (*usability*), keandalan (*reliability*), keamanan (*security*), tingkat kemampuan pemeliharaan (*maintainability*), dan kemampuan portabel (*portability*) (ISO, 2011). Tingkat kemampuan pemeliharaan perangkat lunak tidak dapat diukur secara langsung, tetapi dapat diukur menggunakan prediksi sebuah model berdasarkan seperangkat atribut internal dalam kualitas perangkat lunak (Land, 2002). Dalam tingkat kemampuan pemeliharaan perangkat lunak, terdapat beberapa jenis pengukuran, salah satunya adalah upaya pemeliharaan berdasarkan perubahan (H. Alsolai & Roper, 2020).

Model prediksi yang akurat dari tingkat kemampuan pemeliharaan perangkat lunak telah diselidiki dapat mendukung dan membantu dalam pengambilan keputusan terkait perangkat lunak, efisiensi proses pemeliharaan, membandingkan produktivitas dan biaya di antara berbagai proyek, mengalokasikan sumber daya, dan sebagainya (Elish & Elish, 2009). Model individu atau penggunaan satu

algoritma pembelajaran mesin dalam pembuatan model seperti *general regression neural network* (GRNN) (Thwin & Quah, 2005), *bayesian network* (Van Koten & Gray, 2006), *multivariate adaptive regression splines* (MARS) (Zhou & Leung, 2007), *multilayer perceptron* (MLP) (Dubey dkk., 2012), *mamdani fuzzy logic based* (Ahmed & Al-Jamimi, 2013), *Neuro-GA* (Kumar dkk., 2015), *k-nearest neighbors* (KNN) (H. Alsolai dkk., 2018), *artificial neural network* (ANN), *functional link artificial neural network* (FLANN), *radial basis function neural* (RBFN) *network* (Kumar dkk., 2019) telah diusulkan oleh penelitian sebelumnya untuk memprediksi tingkat kemampuan pemeliharaan perangkat lunak pada sistem berorientasi objek.

Namun menurut H.Alsolai & Roper (2020), dalam enam dari tujuh penelitian sebelumnya, model prediksi tingkat kemampuan pemeliharaan perangkat lunak yang dihasilkan masih belum memenuhi standar akurasi yang ditetapkan dan menjadi tantangan bagi para peneliti. Salah satu cara yang telah terbukti untuk meningkatkan performa adalah dengan menggunakan model ensambel atau gabungan dari beberapa model individu (H. Alsolai dkk., 2018). Selain itu, penggunaan teknik seleksi fitur dapat menjadi tambahan sebagai cara untuk meningkatkan performa dari model (Kumar dkk., 2019). Banyak penelitian yang menggunakan *dataset Quality Evaluation System* (QUES) (Thwin & Quah, 2005) (H. Alsolai dkk., 2018) dan *User Interface Management System* (UIMS) (Van Koten & Gray, 2006) (Zhou & Leung, 2007) (Elish & Elish, 2009) (Kumar dkk., 2019).

Dataset ini ditulis dalam bahasa pemrograman *Classic-Ada* dan terdapat sebelas metrik dalam kedua *dataset* tersebut, sepuluh metrik merupakan independen variabel dan satu metrik merupakan dependen variabel (Li & Henry, 1993). *Dataset* ini banyak digunakan pada penelitian sebelumnya karena bersifat publik, sehingga mudah di replikasi dan memudahkan dalam membandingkan hasilnya dengan penelitian sebelumnya (Kumar dkk., 2019). Namun, *dataset* ini memiliki keterbatasan, di antaranya adalah jumlah data yang relatif kecil dan sudah cukup tua (H. Alsolai & Roper, 2020). Selain itu, bahasa pemrograman dari *dataset* yang digunakan adalah hanya *Classic-Ada*, sedangkan bahasa pemrograman berorientasi objek yang ada saat ini ada banyak, seperti C++, Java, PHP dan sebagainya.

Sehingga validitas data dan kemampuan generalisasi dari model yang dihasilkan perlu dipertanyakan (H. Alsolai dkk., 2018) (Kumar dkk., 2019).

Oleh karena itu, diperlukan penelitian dengan menggunakan *dataset* yang lebih baru, lebih besar, dan menggunakan bahasa pemrograman yang berbeda untuk meningkatkan kemampuan generalisasi dan performa model prediksi tingkat kemampuan pemeliharaan perangkat lunak pada sistem berorientasi objek. Kemudian, teknik seleksi fitur digunakan untuk mengidentifikasi metrik yang paling sesuai dalam *dataset* yang digunakan. Selain itu, model ensambel telah dipertimbangkan untuk digunakan dalam penelitian ini. Pengujian performa dari model akan diukur menggunakan evaluasi pengukuran untuk masalah regresi. Hasil dari pengujian performa tersebut akan di bandingkan dengan model-model yang telah digunakan dalam penelitian terdahulu.

1.2 Rumusan Masalah Penelitian

Berdasarkan latar belakang yang telah diuraikan di atas, maka didapatkan rumusan masalah sebagai berikut:

1. Bagaimana dampak sumber himpunan data yang lebih besar dan berbeda terhadap performa model prediksi tingkat kemampuan pemeliharaan perangkat lunak?
2. Seberapa efisien model ensambel jika digunakan untuk memprediksi upaya pemeliharaan perangkat lunak berdasarkan perubahan?
3. Bagaimana pengaruh teknik seleksi fitur (*feature selection techniques*) terhadap performa model prediksi tingkat kemampuan pemeliharaan perangkat lunak?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis dampak sumber himpunan data yang lebih besar dan berbeda terhadap performa model prediksi tingkat kemampuan pemeliharaan perangkat lunak.
2. Menerapkan model ensambel untuk memprediksi upaya pemeliharaan perangkat lunak berdasarkan perubahan sehingga performa model yang dihasilkan lebih baik.

3. Menganalisis pengaruh penggunaan teknik seleksi fitur (*feature selection techniques*) terhadap performa model yang dihasilkan untuk prediksi tingkat kemampuan pemeliharaan perangkat lunak.

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Hasil dari penelitian ini dapat dijadikan sebagai acuan mengenai perbedaan performa dari penggunaan *dataset* berukuran kecil dan *dataset* berukuran besar.
2. Hasil dari penelitian ini dapat dijadikan sebagai acuan mengenai perbedaan performa dari model ensambel dan model individu
3. Secara umum, dengan prediksi model yang lebih akurat dari penelitian ini dapat digunakan sebagai alat untuk memprediksi upaya pemeliharaan perangkat lunak berdasarkan perubahan.

1.5 Batasan Penelitian

Penelitian ini akan dilakukan dengan beberapa batasan sebagai berikut:

1. Penelitian berfokus pada prediksi tingkat kemampuan pemeliharaan perangkat lunak pada sistem berorientasi objek atau *Object Oriented (OO)*.
2. Penelitian berfokus pada upaya pemeliharaan berdasarkan perubahan (*Change maintenance effort*).
3. Penelitian menggunakan dataset publik yang di dalamnya terdapat beberapa sistem yang ditulis dalam bahasa Java.
4. Penelitian mempertimbangkan teknik pembelajaran mesin menggunakan tujuh model individu dan dua metode ensambel.
5. Penelitian membandingkan hasil antara prediksi model.

1.6 Hipotesis Penelitian

Berikut merupakan hipotesis dari penelitian ini:

1. Sumber himpunan data yang bersifat lebih besar dan berbeda meningkatkan kemampuan generalisasi model prediksi tingkat kemampuan pemeliharaan perangkat lunak.
2. Performa yang dihasilkan dari model ensambel prediksi tingkat kemampuan pemeliharaan perangkat lunak lebih baik dibandingkan dengan menggunakan model individu.

3. Teknik seleksi fitur (*feature selection techniques*) meningkatkan performa model prediksi tingkat kemampuan pemeliharaan perangkat lunak dengan pemilihan fitur yang tepat.

1.7 Struktur Organisasi Skripsi

Sistematika penulisan yang digunakan dalam penelitian ini secara garis besar mengacu pada Pedoman Penulisan Karya Ilmiah Universitas Pendidikan Indonesia Tahun 2019. Adapun untuk strukturnya adalah sebagai berikut:

1. Bab I: Pendahuluan

Pada Bab I Pendahuluan berisi latar belakang penelitian, rumusan masalah penelitian, tujuan penelitian, manfaat penelitian, batasan penelitian, hipotesis penelitian, dan struktur organisasi skripsi. Latar belakang penelitian membahas mengenai penentuan bidang penelitian yang dilakukan dan penentuan topik penelitian berdasarkan masalah yang ada. Rumusan masalah penelitian memaparkan poin-poin pertanyaan penelitian berdasarkan identifikasi masalah-masalah spesifik yang menjadi permasalahan di penelitian sebelumnya yang kemudian dijadikan sebagai tujuan penelitian. Manfaat penelitian membahas gambaran mengenai kontribusi pengetahuan yang dapat diberikan oleh hasil penelitian yang dilakukan. Batasan penelitian membahas mengenai ruang lingkup penelitian yang dilakukan sehingga jelas apa yang menjadi fokus penelitian. Hipotesis penelitian memaparkan jawaban sementara terhadap rumusan masalah yang didapat berdasarkan kajian pustaka.

2. Bab II: Kajian Pustaka

Pada Bab II Kajian Pustaka membahas mengenai tinjauan dari literatur yang terkait dengan penelitian yang dilakukan. Bab ini terdiri dari beberapa subbab yang menjelaskan konsep-konsep penting, teori-teori, metode-metode yang ada (*state of the art*), dan hasil penelitian sebelumnya yang terkait dengan penelitian yang dilakukan.

3. Bab III: Metode Penelitian

Pada Bab III Metode Penelitian membahas mengenai hal-hal yang bersifat prosedural untuk melakukan penelitian, seperti desain penelitian, sumber himpunan data, instrumen penelitian, prosedur penelitian, dan analisis data. Desain penelitian menggambarkan alur penelitian yang dilakukan,

menggunakan kerangka kerja metodologi penelitian yang sesuai. Sumber himpunan data membahas mengenai data yang digunakan dalam penelitian, seperti data apa yang digunakan, berasal dari mana data tersebut, berapa jumlah data yang akan digunakan, dan bagaimana karakteristik data yang digunakan. Instrumen penelitian memaparkan mengenai alat pengumpul data yang dipergunakan dalam penelitian, seperti komputasi yang digunakan dan perangkat lunak yang digunakan dalam penelitian. Prosedur penelitian memaparkan langkah-langkah penelitian yang dilakukan berdasarkan alur penelitian yang dibuat sebelumnya. Analisis data membahas mengenai rumus-rumus dan perangkat lunak untuk analisis statistik yang digunakan untuk evaluasi dalam penelitian.

4. Bab IV: Temuan dan Pembahasan

Pada Bab IV Temuan dan Pembahasan memaparkan hasil yang diperoleh dari proses penelitian yang telah dilakukan, yang berupa jawaban atas pertanyaan penelitian, tujuan penelitian, hipotesis yang diusulkan, dan hasil-hasil lain yang diperoleh dari proses penelitian. Hasil penelitian tersebut dijelaskan secara rinci, dilakukan interpretasi dan dikaitkan dengan teori-teori atau hipotesis untuk menjawab pertanyaan penelitian yang telah dirumuskan sebelumnya.

5. Bab V: Simpulan, Implikasi, dan Rekomendasi

Pada Bab V Simpulan, Implikasi, dan Rekomendasi membahas mengenai kesimpulan yang diperoleh dari hasil analisis temuan penelitian dan interpretasi hasil penelitian. Selain itu, membahas mengenai dampak yang dapat berguna bagi penelitian sejenis dan saran untuk melakukan penelitian selanjutnya.