

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Ketepatan penyelesaian *issue* merupakan bagian penting dari pengendalian proyek perangkat lunak karena memengaruhi ketercapaian target *sprint*, perencanaan rilis, utilisasi sumber daya, dan kualitas koordinasi lintas tim. Dalam praktik pengembangan perangkat lunak, keterlambatan sering baru terlihat setelah suatu *issue* terlalu lama berada pada alur kerja, berulang kali berpindah status, atau terus berada pada *backlog* aktif tanpa penyelesaian yang jelas. Kondisi tersebut menunjukkan bahwa pemantauan proyek tidak cukup dilakukan secara deskriptif, tetapi perlu dilengkapi dengan mekanisme prediktif yang dapat mendeteksi risiko durasi tinggi sejak dini.

Hasil eksplorasi data penelitian menunjukkan bahwa *strict dataset* Day-7 memuat 78.489 *issue*. Dari 15.698 *issue* pada data uji, terdapat 96 *issue* atau sekitar 0,61% yang termasuk dalam kelas risiko keterlambatan berbasis durasi tinggi. Proporsi tersebut menunjukkan bahwa kasus berisiko merupakan kejadian yang relatif jarang dibandingkan keseluruhan data. Namun, kasus tersebut tetap penting untuk dideteksi karena dapat berlangsung dalam durasi yang panjang serta memengaruhi pengendalian *backlog*, penentuan prioritas, dan perencanaan penyelesaian pekerjaan.

Selain ditunjukkan oleh karakteristik data, kebutuhan akan mekanisme *early warning* juga didukung oleh kondisi operasional dalam pelaksanaan proyek perangkat lunak. Manajer proyek dapat menangani kegiatan koordinasi, rapat, pelaporan, pemantauan *backlog*, dan pengambilan keputusan secara bersamaan. Kondisi tersebut menyebabkan manajer proyek tidak selalu dapat meninjau perkembangan setiap *issue* secara terus-menerus maupun mengenali pola risiko pada keseluruhan proyek.

Pelaksanaan *daily stand-up* membantu tim menyampaikan kondisi pekerjaan dan kendala yang sedang dihadapi. Namun, informasi dalam *daily stand-up* bergantung pada laporan anggota tim dan umumnya berfokus pada kendala yang telah disadari pada saat pertemuan. Mekanisme tersebut belum secara otomatis menganalisis pola historis dari seluruh *issue* yang tercatat pada Jira. Oleh karena itu, sistem *early warning* dalam penelitian ini tidak dimaksudkan untuk menggantikan

daily stand-up, tetapi untuk melengkapinya dengan indikator kuantitatif berbasis data historis. Indikator tersebut dapat membantu manajer proyek memprioritaskan *issue* yang perlu diperiksa lebih awal tanpa harus meninjau seluruh *issue* secara manual.

Perkembangan *machine learning* pada manajemen proyek menunjukkan bahwa estimasi durasi, prediksi risiko, dan peramalan kinerja proyek menjadi area yang terus berkembang [1, 2]. Pada ranah *issue tracking*, data Jira juga semakin relevan karena memuat informasi historis mengenai status, prioritas, *assignee*, komentar, dan perubahan alur kerja [3, 4].

Sejumlah penelitian terdahulu telah menunjukkan bahwa data *issue tracker* dapat digunakan untuk memprediksi prioritas, waktu resolusi, dan keterlambatan. Madaraboina dkk. menggunakan pendekatan *multi-target classification* untuk memprediksi prioritas dan waktu resolusi *bug* [5], Özkan dkk. menganalisis prediksi waktu resolusi *bug* pada data Jira [6], sedangkan Qiao dkk. membangun *dataset snapshot* untuk prediksi waktu resolusi *issue open-source software* (OSS) [7].

Penelitian terbaru juga menunjukkan bahwa CatBoost telah digunakan pada tugas prediksi dalam rekayasa perangkat lunak. De Rosa dan Liguori [8] membandingkan Random Forest, XGBoost, dan CatBoost untuk memprediksi *residual defect* pada sistem Python dengan memanfaatkan metrik produk, proses, statistik, dan aktivitas pengembang. Hasil penelitian tersebut menunjukkan bahwa model *supervised* berbasis metrik, termasuk CatBoost, mampu menghasilkan *recall* yang tinggi dan bahwa metrik proses memberikan informasi prediktif yang penting. Meskipun target dan unit analisisnya berbeda, temuan tersebut mendukung penggunaan CatBoost dan fitur proses pada prediksi berbasis data rekayasa perangkat lunak.

Meskipun demikian, dua permasalahan utama masih perlu dikaji. Pertama, penelitian terdahulu belum banyak mengintegrasikan metadata *issue*, fitur proses dari riwayat perubahan status, dan histori *assignee* dalam satu rancangan *early warning* yang memperbarui prediksi pada beberapa horizon observasi. Kedua, kontribusi fitur proses perlu diuji terhadap *baseline* pada sumber data, target, pembagian temporal, parameter dasar, dan prosedur evaluasi yang sama. Dengan demikian, perubahan kinerja dapat dikaitkan dengan perbedaan kelompok fitur dan bukan dengan perbedaan sumber data atau skema pengujian.

Data Jira memiliki karakteristik tabular campuran yang terdiri atas fitur kategorikal, seperti proyek, tipe *issue*, prioritas, dan status, serta fitur numerik,

seperti jumlah komentar, jumlah transisi, dan durasi historis. CatBoost dipilih karena dapat memproses fitur kategorikal dan numerik dalam satu model serta mendukung pembobotan kelas dan interpretasi kontribusi fitur. Penggunaan algoritma yang sama pada seluruh konfigurasi juga memungkinkan perbandingan difokuskan pada perbedaan kelompok fitur, bukan pada perbedaan algoritma.

Atas dasar tersebut, diperlukan model prediksi risiko durasi tinggi yang memadukan fitur Jira standar, fitur proses berbasis *event log*, dan statistik historis *assignee*, lalu divalidasi terhadap pendekatan terdahulu. Selain menghasilkan skor risiko, model prediktif juga dilengkapi penjelasan lokal untuk menampilkan faktor utama yang memengaruhi skor kepada pengguna [9, 10].

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, rumusan masalah dirumuskan sebagai berikut.

1. Bagaimana membangun model *early warning* dinamis berbasis CatBoost untuk memprediksi risiko keterlambatan *issue* yang dioperasikan sebagai risiko durasi tinggi pada horizon Day-1, Day-7, dan Day-30?
2. Bagaimana kinerja model dan kontribusi fitur proses pada target risiko durasi tinggi serta estimasi waktu tersisa berdasarkan evaluasi temporal?

1.3 Batasan Permasalahan

Batasan permasalahan dalam penelitian ini ditetapkan sebagai berikut.

1. Penelitian menggunakan CatBoost sebagai algoritma klasifikasi untuk memprediksi risiko keterlambatan *issue* berdasarkan data Jira. Penelitian tidak melakukan perbandingan eksperimental antarkeluarga algoritma sebagai fokus utama.
2. Data penelitian dibatasi pada The Public Jira Dataset dalam format *dump* MongoDB. Hasil penelitian mengikuti karakteristik dan kelengkapan atribut yang tersedia pada *dataset* tersebut.
3. Mekanisme *early warning* dinamis dibatasi pada horizon Day-1, Day-7, dan Day-30. Day-1 mewakili indikasi awal, Day-7 mewakili pemeriksaan setelah

satu minggu, dan Day-30 mewakili eskalasi setelah satu bulan. Horizon Day-14 dan Day-21 tidak dievaluasi sehingga hasil penelitian tidak digunakan untuk menyimpulkan kinerja model pada kedua horizon tersebut. Istilah dinamis merujuk pada pembaruan representasi kondisi *issue* pada setiap horizon, bukan pada pelatihan ulang model secara terus-menerus.

4. Risiko keterlambatan direpresentasikan oleh durasi penyelesaian yang melebihi kuartil ke-75 data latih. Definisi tersebut merupakan definisi operasional penelitian dan bukan batas SLA atau *due date* resmi.
5. Selain target risiko keterlambatan berbasis durasi tinggi, penelitian mengevaluasi target estimasi waktu tersisa setelah Day-7 dan Day-30. Pada Day-7, target menunjukkan apakah sisa waktu penyelesaian melebihi 30 hari atau 90 hari. Pada Day-30, target menunjukkan apakah sisa waktu penyelesaian melebihi 30 hari, 60 hari, atau 90 hari. Seluruh target tersebut berupa klasifikasi biner dan bukan prediksi jumlah hari secara kontinu.
6. *Process mining* digunakan untuk mentransformasikan riwayat perubahan status Jira menjadi *event log* dan fitur proses. Penelitian tidak mencakup pembangunan BPMN, Petri net, analisis kesesuaian proses, atau *process discovery* penuh.
7. Evaluasi model dibatasi pada pembagian data secara temporal, metrik ROC-AUC, *Average Precision*, *precision*, *recall*, *F1-score*, *accuracy*, *confusion matrix*, serta uji McNemar pada pasangan model yang sesuai.
8. Kontribusi fitur digunakan untuk menjelaskan keputusan prediktif model. Hasil tersebut tidak digunakan untuk menyimpulkan hubungan sebab-akibat maupun menilai kinerja personal seorang *assignee*.
9. Antarmuka yang dibangun hanya digunakan sebagai prototipe demonstrasi dan tidak dievaluasi melalui pengujian kegunaan atau pendekatan *user-centered design*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan penelitian ini adalah sebagai berikut.

1. Membangun model *early warning* dinamis berbasis CatBoost untuk memprediksi risiko keterlambatan *issue* yang dioperasionalkan sebagai risiko durasi tinggi pada horizon Day-1, Day-7, dan Day-30.
2. Mengevaluasi kinerja model dan kontribusi fitur proses pada target risiko durasi tinggi serta estimasi waktu tersisa menggunakan pembagian data temporal.

1.5 Manfaat Penelitian

Manfaat yang diharapkan adalah sebagai berikut:

1. Bagi praktisi, hasil penelitian dapat menjadi dasar sistem *early warning* untuk mendeteksi *issue* yang berisiko berdurasi panjang.
2. Bagi organisasi, hasil penelitian dapat membantu penentuan prioritas penanganan, pengawasan *backlog*, dan identifikasi pekerjaan yang perlu diekskalasi.
3. Bagi akademisi, hasil penelitian menambah kajian empiris mengenai pemanfaatan data Jira, fitur proses, statistik *assignee*, dan model prediktif pada manajemen proyek perangkat lunak.
4. Bagi pengembangan metodologi, hasil penelitian menunjukkan perbandingan antara fitur statis, fitur proses, fitur histori *assignee*, dan pendekatan penjelasan lokal berbasis kontribusi fitur.

1.6 Sistematika Penulisan

Sistematika penulisan laporan disusun sebagai berikut:

1. Bab 1 PENDAHULUAN

Bab ini berisi latar belakang masalah, rumusan masalah, batasan permasalahan, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

2. Bab 2 LANDASAN TEORI

Bab ini memuat penelitian terdahulu dan landasan teori mengenai Jira, prediksi waktu resolusi, *early warning*, *process mining*, PM4Py, CatBoost, metrik evaluasi, dan *explainability*.

3. **Bab 3 METODOLOGI PENELITIAN**

Bab ini menjelaskan sumber data, pembersihan data, pembentukan label, pembentukan fitur, perancangan model, pembagian data, evaluasi, dan rancangan inferensi.

4. **Bab 4 HASIL DAN PEMBAHASAN**

Bab ini menyajikan hasil eksperimen, perbandingan model, analisis horizon prediksi, interpretasi faktor, uji statistik, dan pembahasan hasil penelitian.

5. **Bab 5 PENUTUP**

Bab ini memuat simpulan, keterbatasan, dan saran pengembangan berikutnya.

