

# BAB 1. PENDAHULUAN

## 1.1. Latar Belakang

*Software delivery* yang sesuai dengan kebutuhan yang telah ditentukan merupakan salah satu tujuan utama dalam pengembangan perangkat lunak komersial [1]. Untuk mencapai kualitas yang sesuai standar, diperlukan serangkaian pengujian terstruktur, dengan *User Acceptance Testing* (UAT) sebagai tahap krusial terakhir sebelum implementasi [2].

UAT memegang peranan penting sebagai penghubung antara pengembang dan pengguna akhir, memastikan perangkat lunak yang dibangun sesuai dengan kebutuhan pengguna dan tujuan bisnis [3]. Tantangan utama dalam UAT adalah memastikan cakupan pengujian yang luas dan mencakup beragam skenario penggunaan nyata, yang sering kali membutuhkan biaya, waktu, dan sumber daya yang signifikan.

*Crowdsourcing* menawarkan solusi potensial untuk tantangan tersebut dengan melibatkan banyak pengguna dari berbagai latar belakang dalam proses pengujian, sehingga perusahaan mendapatkan cakupan pengujian yang lebih luas dan realistis dengan alokasi anggaran yang lebih efisien [6]. Namun, pendekatan ini menghadirkan tantangan baru berupa tidak tersaringnya penguji (*workers*) yang layak dari yang tidak layak.

Ketidak layakan *workers* dalam *crowdsourcing* membuka celah kecacatan pengujian, di mana penguji memiliki probabilitas untuk tidak memenuhi standar kebutuhan pengujian. Menurut penelitian Kazai et al. (2011), "ketidak layakan *workers* dalam *crowdsourcing* dapat menyebabkan hasil yang tidak konsisten akibat perbedaan pengalaman, perhatian, dan pemahaman terhadap persyaratan tugas" [7]. Dalam konteks ini, tidak semua penguji mampu atau bersedia mengikuti standar. Ketidak layakan ini dipengaruhi oleh motivasi, latar belakang, keterampilan, dan perilaku *workers*, dengan risiko adanya *workers* yang tidak jujur, asal-asalan, atau malas [8].

Konsekuensinya, ketidak layakan *workers* berpotensi memperlambat waktu

pengujian, mengganggu proses distribusi perangkat lunak, dan menghasilkan hasil uji yang bias [7]. Dibutuhkan mekanisme klasifikasi *workers* yang efektif untuk memastikan proses pengujian berjalan optimal dan menghasilkan data yang andal.

Algoritma M-X yang dikembangkan oleh D. Dang, Y. Liu, X. Zang, dan S. Huang [9]. menawarkan pendekatan menjanjikan untuk menilai kualitas *workers* berdasarkan konsistensi jawaban. Keunggulan algoritma ini terletak pada kemampuannya mengidentifikasi *workers* yang konsisten tanpa bergantung pada jawaban baku, sehingga dapat diterapkan pada skenario pengujian yang bersifat subjektif maupun eksploratif seperti pengujian *usability*, validasi fitur, atau evaluasi berbasis skenario realistis. Namun, tingkat akurasi algoritma ini dalam konteks UAT berbasis *crowdsourcing* belum dikaji secara komprehensif, dan belum jelas sejauh mana algoritma ini dapat memisahkan penguji yang layak untuk pengujian dari yang tidak layak untuk memastikan konsistensi kualitas pengujian.

Penelitian ini mengusulkan pendekatan empiris untuk mengevaluasi efektivitas algoritma M-X dalam klasifikasi kelayakan *workers* pada konteks UAT berbasis *crowdsourcing*. Melalui eksperimen kuantitatif, penelitian ini akan mengukur keakuratan algoritma dalam mengidentifikasi pola konsistensi jawaban di antara penguji, serta kemampuannya beradaptasi dengan lingkungan pengujian UAT berbasis *crowdsourcing*.

Untuk mendukung eksperimen tersebut, sistem simulasi berbasis arsitektur *client-server* akan dikembangkan dengan GraphQL sebagai antarmuka pengiriman data dan MongoDB sebagai basis data. Pemilihan teknologi ini mempertimbangkan kebutuhan fleksibilitas dan skalabilitas sistem [10], kemampuan GraphQL dalam mengelola struktur data kompleks secara efisien [11], serta keunggulan MongoDB dalam menangani data tidak terstruktur dengan performa tinggi [12]. Penelitian ini diharapkan menghasilkan evaluasi komprehensif terhadap efektivitas algoritma M-X sebagai mekanisme kontrol kualitas dalam lingkungan pengujian berbasis *crowdsourcing*.

## 1.2. Perumusan Masalah

Setelah memahami latar belakang dan konteks penelitian, berikut adalah rumusan masalah yang menjadi fokus penelitian:

1. Bagaimana tingkat akurasi algoritma M-X dalam mengklasifikasikan pengujian dalam lingkungan UAT berbasis *crowdsourcing*?
2. Sejauh mana algoritma M-X dapat membantu memisahkan kandidat pengujian yang layak dengan yang tidak layak untuk melakukan UAT berbasis *crowdsourcing*?

Melalui kedua rumusan masalah ini, penelitian bertujuan untuk memberikan pemahaman mendalam tentang kesesuaian algoritma M-X untuk seleksi pengujian yang andal dalam konteks modern UAT berbasis *crowdsourcing*.

## 1.3. Tujuan Penelitian

Tujuan utama dari penelitian ini adalah untuk menjawab rumusan masalah dengan mengimplementasikan penelitian kuantitatif menggunakan sistem simulasi yang dirancang secara metodologis dan menganalisis data metrik kinerja algoritma secara sistematis. Hasil dari penelitian ini diharapkan dapat memberikan landasan empiris bagi pengembangan sistem seleksi pengujian yang lebih adaptif dan akurat dalam lingkungan UAT *crowdsourcing* yang lebih dinamis.

Tabel 1.1. Tabel Keterkaitan Antara Tujuan, Pengujian dan Kesimpulan.

No.	Tujuan	Pengujian	Kesimpulan
1.	Mengevaluasi efektivitas algoritma M-X dalam klasifikasi kelayakan pengujian pada pengujian UAT berbasis <i>crowdsourcing</i> .	Mengadakan simulasi terkontrol untuk membawa algoritma M-X ke ranah seleksi pengujian untuk UAT berbasis <i>crowdsourcing</i> .	Analisis dilakukan untuk menilai sejauh mana algoritma M-X mampu menyaring pengujian yang konsisten.

#### 1.4. Batasan Masalah

Untuk mencegah meluasnya cakupan dari masalah ini yang berpotensi memudarkan fokus penelitian, maka dibuatlah batasan-batasan permasalahan untuk penelitian dan perancangan sebagai berikut:

1. Penelitian ini mencakup kontrol kualitas penguji dalam konteks UAT *crowdsourcing*, dan tidak mencakup pengujian pada tahap pengujian perangkat lunak lainnya, seperti contoh *Unit Testing* atau *Integration Testing*.
2. Penelitian ini akan mempertimbangkan keterbatasan waktu, biaya, dan sumber daya dalam proses kontrol kualitas, sehingga pendekatan yang dirancang akan berorientasi pada solusi yang efisien dan praktis.
3. Fokus utamanya adalah pada evaluasi terhadap metode untuk memisahkan penguji yang layak dari yang tidak layak untuk pengujian, tanpa melibatkan aspek teknis implementasi UAT secara keseluruhan.
4. Simulasi hanya dilakukan pada skenario pilihan ganda, sehingga algoritma M-X tidak diuji untuk soal tipe lain seperti isian atau esai.
5. Penentuan *ground truth* dilakukan secara manual dan terbatas pada skenario pengujian tertentu, sehingga validitas absolut tidak dijamin secara menyeluruh.
6. Evaluasi hanya dilakukan terhadap konsistensi antar jawaban (*peer consistency*), bukan terhadap efektivitas hasil pengujian terhadap fungsionalitas perangkat lunak secara menyeluruh.
7. Sistem yang dikembangkan tidak diuji pada lingkungan produksi nyata, tetapi hanya pada simulasi internal yang dikendalikan.
8. Jumlah penguji dalam simulasi dibatasi maksimal 24 orang, yang mungkin belum mencerminkan kompleksitas *crowd* dalam skenario industri sebenarnya.
9. Faktor eksternal seperti motivasi, kelelahan, atau kontekstual lingkungan pengujian tidak dikendalikan secara ketat, sehingga dapat mempengaruhi hasil klasifikasi kualitas penguji.
10. Relawan penelitian akan diambil mayoritas dari lingkup Telkom

University karena keterbatasan sumber daya dan waktu.

## **1.5. Metode Penelitian**

Penelitian ini menggunakan metode penelitian kuantitatif, lebih tepatnya simulasi kuantitatif, dengan variabel untuk mengamati dan mengukur pengaruh algoritma M-X dalam proses klasifikasi kualitas penguji. Eksperimen dilakukan melalui sistem berbasis web yang dirancang khusus untuk mengimplementasikan algoritma tersebut dalam konteks *User Acceptance Testing* (UAT) berbasis *crowdsourcing*. Data yang dihasilkan dari simulasi bersifat numerik dan dianalisis untuk mengukur efektivitas klasifikasi menggunakan metrik seperti akurasi, presisi, dan *recall*.

### **1.5.1. Kajian Pustaka**

Pada tahap ini dilakukan studi literatur dan pengumpulan referensi jurnal yang sesuai dengan topik yang diangkat pada penelitian ini. Studi literatur ini mencakup *Crowdsourcing Software Testing*, *Usability Testing*, kontrol kualitas *workers*, serta penggunaan algoritma untuk klasifikasi *workers* berdasarkan keterampilan dan kualitas. Informasi yang didapat akan digunakan sebagai landasan teori dan acuan dalam mengembangkan metode kontrol kualitas *workers* pada penelitian ini.

### **1.5.2. Rancangan Penelitian**

Kegiatan rancangan penelitian ini akan terdiri dari perancangan sistematika penyelesaian masalah, perancangan sistem, baik itu desain arsitektur, maupun desain model, dan perancangan skema pengujian penelitian. Penelitian ini berfokus pada implementasi metode kontrol kualitas *workers* pada *User Acceptance Testing* berbasis *Crowdsourcing*. Penelitian ini akan mengimplementasikan algoritma M-X di dalam ranah *platform* UAT dan melihat hasil dari implementasi algoritma tersebut.

### **1.5.3. Pengembangan Sistem**

Dalam rangka mewujudkan keberhasilan dari penelitian ini, maka dibangun sistem berbasis web yang dapat menyeleksi kandidat *workers* untuk

melakukan *User Acceptance Testing* berbasis *crowdsourcing* dengan mengimplementasikan algoritma M-X. Sistem ini akan mencakup fitur pengelolaan data *workers*, klasifikasi berdasarkan kualitas, dan rekomendasi *workers* terbaik untuk setiap tugas pengujian. Sistem akan dibangun menggunakan kerangka *framework Next.js* untuk antarmuka pengguna dan Nest.js sebagai *backend*, serta menggunakan *database* non-relasional seperti mongoDB untuk manajemen data.

#### **1.5.4. Simulasi dan Uji Coba**

Penelitian ini menggunakan metode pengumpulan data kuantitatif dan kualitatif. Data kuantitatif didapatkan melalui hasil uji coba sistem oleh pengguna sebagai kandidat penguji. Uji coba yang dijalankan akan menerapkan algoritma M-X terhadap proses kontrol kualitas di sistem. Data kualitatif diperoleh dari hasil simulasi proses pengujian UAT oleh para penguji yang lolos uji coba. Data ini akan dianalisis untuk mengidentifikasi pola perilaku dan karakteristik *workers*.

#### **1.5.5. Analisis Hasil Simulasi**

Analisis hasil simulasi dilakukan menggunakan pendekatan evaluasi berbasis *Confusion Matrix* untuk mengukur performa algoritma M-X dalam mengklasifikasikan penguji. Metrik evaluasi utama yang dihitung meliputi akurasi (*accuracy*), presisi (*precision*), *recall*, *True Negative Rate* (TNR), dan *F1 Score* yang masing-masing memberikan perspektif berbeda terhadap kemampuan klasifikasi algoritma.

#### **1.5.6. Penyusunan Laporan / Buku TA**

Semua hasil penelitian, mulai dari kajian pustaka, pengumpulan data, analisis, pengembangan sistem, hingga pengujian, akan didokumentasikan dalam bentuk laporan akhir. Laporan ini akan mencakup kesimpulan dan rekomendasi untuk penelitian lebih lanjut.