**Latar Belakang**

Pengembangan perangkat lunak melibatkan berbagai perusahaan dan individu dengan latar belakang pengetahuan dan keahlian yang beragam [1]. Meskipun aplikasi perangkat lunak telah banyak berhasil diterapkan di berbagai instansi, tingkat kegagalannya masih cukup tinggi [2]. Berdasarkan laporan *Standish Group*, hanya 22% proyek pengembangan perangkat lunak di Asia yang berhasil, sementara 58% menghadapi masalah, dan 20% gagal total [3]. Di Indonesia sendiri, tingkat keberhasilan implementasi proyek perangkat lunak masih rendah, hanya mencapai 50% [4]. Kegagalan proyek ini umumnya disebabkan oleh kelemahan dalam proses pengujian [5].

Dalam sebuah proyek, mendeteksi dan memperbaiki *bug* adalah kegiatan yang memakan banyak waktu dan biaya [6]. *Bug* dapat dicegah melalui proses pengujian perangkat lunak [7]. Salah satu elemen penting dalam pengujian adalah *test case*, yang merupakan serangkaian instruksi yang dirancang untuk menemukan kesalahan yang dapat menyebabkan kegagalan sistem [8]. Namun, pembuatan *test case* secara manual adalah proses yang memerlukan waktu dan biaya yang besar [9].

Penelitian sebelumnya yang berjudul "*Designing a Test Case Catalog based on Document Review for Optimizing Test Case Development Process*" menggunakan pendekatan document review untuk perancangan Katalog Test Case [6]. Pada tahap perancangannya, hanya menggunakan alat wordcount yang menghasilkan sembilan kata fungsi yang sering muncul dari dua dokumen kebutuhan perangkat lunak dan tidak jelas ada berapa frekuensi kata-kata fungsinya. Pendekatan ini terbatas karena tidak dapat menganalisis kata fungsi lain yang sering muncul atau mungkin penting, serta tidak melibatkan penggunaan *threshold* dari distribusi frekusensi fitur untuk mengeliminasi kata-kata yang kurang informatif atau jarang digunakan dalam dokumen kebutuhan perangkat lunak. Selain itu, penelitian tersebut tidak menerapkan teknik *data science* seperti *text mining*, yang seharusnya dapat memberikan analisis yang lebih mendalam terhadap kemunculan kata-kata fungsi dalam dokumen kebutuhan perangkat lunak [6].

Idealnya, dokumen Kebutuhan Pengembangan Perangkat Lunak (KPPL) yang memuat requirement sudah diselesaikan sebelum proses pengembangan dan pengujian dimulai. Namun, dalam proyek pengembangan sistem kampus XYZ, terdapat aplikasi yang sudah di deploy dan membutuhkan pengujian karena adanya penambahan fitur baru. Proyek tersebut belum memiliki dokumen KPPL, dan dokumen pengujiannya masih dalam tahap awal, dengan jumlah *test case* yang masih sedikit. Berdasarkan wawancara, mereka belum memiliki repositori *test case*, dan tantangan terbesar yang dihadapi adalah tenggat waktu pengujian, perubahan requirement, serta kompleksitas sistem.

Untuk mengatasi masalah ini, penelitian ini bertujuan untuk merancang Katalog Test Case yang menggunakan prinsip *test case reuse* dengan pendekatan *text mining* untuk menyeleksi *test case* yang relevan. Dengan menggunakan test case yang sudah ada dan relevan dibandingkan membuat dari awal, penelitian ini diharapkan dapat memberikan keuntungan dalam penghematan waktu dan biaya, serta membantu mengatasi kendala proses pengujian perangkat lunak. Untuk menyeleksi *test case* yang relevan, diperlukan penelusuran fungsi atau fitur yang sering digunakan di berbagai aplikasi. Penggunaan kembali *test case* ini harus didasarkan pada fungsi yang relevan, sesuai dengan penelitian terdahulu [10] [11]. *Test case* yang dapat digunakan kembali ini diambil dari dokumen kebutuhan perangkat lunak yang digunakan dalam proyek sistem kampus XYZ. Dokumen tersebut berisi banyak fungsi atau fitur, yang perlu diekstrak nilainya. Proses ekstraksi fitur ini sangat penting dalam pemrosesan dokumen [12]. Salah satu teknik ekstraksi fitur yaitu *Bag of Words* (BoW), yang dapat digunakan untuk menganalisis kata-kata yang relevan dalam konteks topik dan isu dalam berita-berita di Bangladesh berdasarkan penelitian sebelumnya [13]. Teknik ekstraksi fitur ini bertujuan untuk menganalisis dan menyeleksi fitur yang paling relevan dengan menggunakan *threshold* berdasarkan nilai persentil ke-70 dari distribusi frekuensi fitur. Pemilihan persentil ke-70 didasarkan pada asumsi bahwa fitur-fitur yang frekuensinya berada di atas nilai ini lebih sering muncul dan dianggap lebih relevan untuk digunakan kembali di berbagai dokumen kebutuhan. Sebaliknya, fitur yang berada di bawah persentil ke-70 dianggap kurang signifikan karena kemunculannya yang rendah, sehingga lebih mungkin untuk dieliminasi. Dengan demikian, analisis akan lebih fokus pada fitur-fitur yang memiliki potensi lebih tinggi untuk digunakan kembali dalam berbagai konteks sistem. Kumpulan fitur yang telah diseleksi kemudian diolah lebih lanjut menjadi *test case*. Setelah itu, *test case* diseleksi, dihasilkan, diakumulasi, dan disusun menjadi Katalog Test Case yang dapat membantu pengujian proyek aplikasi sistem kampus XYZ.

**Topik dan Batasannya**

Berdasarkan latar belakang diatas, penelitian ini mengangkat beberapa permasalahan yang dikaji: 1) Bagaimana cara menemukan fungsi yang relevan menggunakan *text mining* untuk merancang Katalog Test Case dari dokumen KPPL agar dapat menghasilkan *test case* yang relevan untuk membantu pengujian pada aplikasi yang sudah di deploy?; 2) Bagaimana efektivitas penggunaan Katalog Test Case?

Berdasarkan rumusan masalah yang telah dijelaskan, maka batasan masalah dari penelitian ini adalah: 1) Penelitian ini dilakukan pada proyek pengembangan perangkat lunak system kampus XYZ karena terdapat aplikasi asrama yang sudah di deploy dan memerlukan pengujian tetapi belum memiliki dokumen KPPL dan dokumen pengujian yang masih tahap awal, dengan test case yang masih sedikit.; 2) Dokumen Kebutuhan Pengembangan Perangkat Lunak (KPPL) yang dipakai adalah dokumen kebutuhan perangkat lunak travel, kesehatan mental, booking tiket, e-learning, sewa sepeda, mobile fakultas, web faktultas, sewa kost, Transkip Aktivitas Kemahasiswaan (TAK), Penerimaan Mahasiswa Baru (PMB), dan kehadiran karyawan.

**Tujuan**

Penelitian ini bertujuan untuk merancang Katalog Test Case bagi tim QA berdasarkan kesamaan fungsi atau fungsi yang relevan menggunakan teknik *text mining*. Katalog Test Case ini berisi test case yang dapat digunakan kembali untuk proyek-proyek lain. Selain itu, penelitian ini juga menguji efektivitas Katalog Test Case dengan menerapkannya pada pengujian proyek lain. Dari hasil pengujian tersebut, diperoleh gambaran tentang dampak nyata penggunaan Katalog Test Case terhadap waktu pengujian, yang pada akhirnya juga memengaruhi biaya yang dikeluarkan.