

pengajar untuk menilai seberapa sesuai jawaban tersebut. Teknologi dapat berupa aplikasi yang terpasang pada sebuah perangkat tertentu (aplikasi *mobile* atau *desktop*) atau berupa aplikasi *online* yang dapat diakses melalui internet (*website*). Solusi ini tentunya akan memberikan penilaian lebih objektif terhadap jawaban esai yang dievaluasi.

Aplikasi ini kemudian juga harus dapat diakses dimanapun dan kapanpun agar penggunaan aplikasi tidak terbatas di satu perangkat saja. Namun, pengembangan aplikasi *mobile* perlu disesuaikan untuk setiap *platform* yang akan semakin membatasi dimana aplikasi dapat digunakan [6]. Maka dari itu, aplikasi berbasis *website* adalah solusi terbaik untuk permasalahan ini karena aplikasi akan dapat digunakan kapan saja dan dimana saja, tanpa membatasi perangkat apa yang dapat mengakses aplikasi tersebut.

Aplikasi akan dikembangkan mengikuti siklus hidup perangkat lunak *Test-Driven Development* (TDD) yang merupakan salah satu *Software Development Life-Cycle* (SDLC) *Agile* [7]. Praktik TDD dimulai dengan memikirkan cara menguji sebuah fungsi. Programmer kemudian menulis kode yang dapat memenuhi pengujian [8]. Dibanding dengan *eXtreme Programming* (XP), dan *Personal eXtreme Programming* (PXP), yang juga termasuk dalam SDLC *Agile*, TDD terbukti dapat menghasilkan aplikasi yang sangat kecil kemungkinannya untuk memiliki *bug* dan *error* karena pada setiap fungsi yang akan dikembangkan perlu dibuat syarat keberhasilan pengujian terlebih dahulu, kemudian jika kode yang dituliskan berhasil memenuhi syarat keberhasilan pengujian tersebut, pengembang dapat melanjutkan pengembangan fungsi selanjutnya [9].

Pengujian aplikasi / hasil program akan dilakukan pada aplikasi yang dikembangkan untuk memastikan kelayakannya menggunakan tiga metode pengujian, yaitu *Integration Testing*, *System Usability Scale* (SUS) dan *Black Box*. *Integration Testing* dilakukan untuk memastikan bahwa setiap fungsi yang dibangun pada aplikasi telah memenuhi syarat kebutuhannya [10]. SUS memiliki tingkat keabsahan dan keandalan tinggi untuk mengukur tingkat kemudahan perangkat lunak saat