

Génération et Évaluation de Feedback pour les Plateformes d'Apprentissage de Programmation en Ligne

MOCAH - Pedagog.ia

Stage de Master - 2025

1 Contexte et Motivation

Ce stage s'inscrit dans le cadre d'une thèse en cours visant à concevoir une politique optimale de feedback pour les apprenants débutants en Python utilisant la plateforme AlgoPython. Cette plateforme interactive d'apprentissage de la programmation Python présente plusieurs caractéristiques notables. Elle permet une catégorisation des exercices par concepts, tels que les boucles, les fonctions et les structures conditionnelles. De plus, elle intègre un système de soumissions itératives, permettant aux apprenants de soumettre leurs solutions plusieurs fois et de recevoir un feedback personnalisé sur leurs erreurs de programmation.

Le mécanisme de feedback actuel de la plateforme AlgoPython présente certaines limitations. Il se concentre fortement sur les erreurs de syntaxe, fournissant des messages d'erreur explicites pour ces types de problèmes. Cependant, pour les erreurs sémantiques, le feedback est souvent générique et manque de granularité, ce qui limite son impact pédagogique. Les indications fournies ne sont pas toujours suffisamment détaillées pour aider efficacement les apprenants à comprendre et corriger leurs erreurs.

Pour illustrer la problématique, considérons l'exemple suivant :

```
for i in range(3):  
    print("Hello, world!")
```

Le comportement attendu est que la boucle devrait itérer 5 fois au lieu de 3 fois. Différents niveaux de feedback peuvent être envisagés pour aider l'apprenant à corriger cette erreur. Un feedback général pourrait simplement indiquer que le nombre d'itérations est incorrect et suggérer de vérifier la boucle. Un feedback suggestif pourrait encourager l'apprenant à revoir le nombre d'itérations attendu. Un feedback explicite préciserait que la boucle s'exécute 3 fois alors qu'elle devrait s'exécuter 5 fois. Enfin, un feedback technique pourrait orienter l'apprenant vers la vérification de l'initialisation et de la condition de la boucle.

2 Objectifs du Stage

2.1 Objectifs Principaux

L'un des objectifs principaux de ce stage est l'amélioration des critères de feedback. Cela inclut l'évaluation des critères existants, le développement de nouveaux critères pertinents et leur validation avec des experts pédagogiques. Un autre objectif majeur est le développement du système de génération de feedback. Ce système intégrera un mécanisme RAG (Retrieval-Augmented Generation) et un raisonnement "chain-of-thought", tout en s'adaptant aux spécificités de la plateforme. Enfin, l'évaluation de la qualité du feedback généré sera également une priorité. Des métriques quantitatives seront développées, des évaluations qualitatives seront mises en place, et des tests utilisateurs seront réalisés avec des apprenants.

2.2 Livrables Attendus

À l'issue de ce stage, plusieurs livrables sont attendus. Un système de génération de feedback fonctionnel sera développé et intégré à la plateforme. Une documentation technique complète sera rédigée pour détailler les aspects techniques du système. Un rapport d'évaluation des performances sera produit pour analyser l'efficacité du feedback généré. Enfin, des recommandations pour l'amélioration continue du système seront formulées.

3 Méthodologie Proposée

3.1 Phase 1 : Analyse et Préparation

La première phase du projet consistera en une étude approfondie de la littérature existante sur les systèmes de feedback en programmation. Cette revue de la littérature permettra d'identifier les meilleures pratiques et les approches innovantes dans le domaine du feedback automatisé. Nous examinerons les travaux antérieurs sur les systèmes de feedback pour la programmation, en mettant l'accent sur les méthodes qui ont démontré leur efficacité pédagogique. Cette analyse nous aidera à comprendre les forces et les faiblesses des systèmes existants et à identifier les opportunités d'amélioration.

Parallèlement, une analyse des patterns d'erreurs communs chez les apprenants débutants pourra être réalisée. Cette analyse sera basée sur des données collectées à partir de la plateforme, ainsi que sur des études de cas et des observations directes. Nous identifierons les types d'erreurs les plus fréquents, tels que les erreurs de syntaxe, les erreurs sémantiques et les erreurs logiques. Cette compréhension approfondie des erreurs courantes nous permettra de définir des critères de qualité pour le feedback. Ces critères viseront à garantir que le feedback soit à la fois pertinent et pédagogiquement efficace, en fournissant des indications claires et constructives pour aider les apprenants à corriger leurs erreurs et à améliorer leurs compétences en programmation.

3.2 Phase 2 : Développement

La deuxième phase sera consacrée à l'implémentation du système RAG (Retrieval-Augmented Generation) et à l'intégration des modèles de langage nécessaires. Le système RAG combinera des techniques de récupération d'informations et de génération de texte pour produire un feedback personnalisé et contextuellement pertinent. Nous utiliserons des modèles de langage avancés, tels que les modèles de type Transformer, pour générer des messages de feedback cohérents et précis.

Des mécanismes d'adaptation au contexte seront développés pour s'assurer que le feedback généré soit toujours pertinent et adapté aux besoins spécifiques de chaque apprenant. Ces mécanismes prendront en compte divers facteurs contextuels, tels que le niveau de compétence de l'apprenant, le type d'exercice et les erreurs spécifiques commises. Par exemple, pour un apprenant débutant, le feedback pourrait être plus détaillé et explicatif, tandis que pour un apprenant plus avancé, le feedback pourrait être plus concis et technique.

Nous intégrerons également un raisonnement "chain-of-thought" dans le système de génération de feedback. Cette approche permettra de décomposer le processus de résolution de problèmes en étapes logiques, fournissant ainsi un feedback structuré et facile à suivre pour les apprenants. Le raisonnement "chain-of-thought" aidera les apprenants à comprendre non seulement ce qui ne va pas dans leur code, mais aussi pourquoi et comment ils peuvent le corriger.

3.3 Phase 3 : Évaluation et Validation

La troisième phase sera dédiée aux tests et à l'évaluation du système. Des tests seront réalisés avec des enseignants en programmation pour valider la pertinence pédagogique du feedback. Ces enseignants évalueront la clarté, la précision et l'utilité des messages de feedback générés par

le système. Leurs retours seront essentiels pour affiner les critères de qualité et améliorer la pertinence pédagogique du feedback.

Des évaluations seront également menées avec des étudiants pour mesurer l'impact du feedback sur leur apprentissage. Nous utiliserons des métriques quantitatives, telles que le taux de correction des erreurs et le temps nécessaire pour résoudre les exercices, pour évaluer l'efficacité du feedback. Des évaluations qualitatives, telles que des enquêtes de satisfaction et des entretiens avec les apprenants, seront également réalisées pour recueillir leurs impressions et suggestions.

En fonction des retours obtenus, des ajustements seront effectués pour améliorer continuellement le système. Ces ajustements pourront inclure des modifications des modèles de langage, des améliorations des mécanismes d'adaptation au contexte et des ajustements des critères de qualité pour le feedback. L'objectif sera de créer un système de feedback dynamique et évolutif, capable de s'adapter aux besoins changeants des apprenants et de contribuer de manière significative à leur apprentissage de la programmation.

4 Compétences Requises

Pour mener à bien ce stage, plusieurs compétences sont requises. Une solide connaissance en Python est essentielle, ainsi qu'une familiarité avec les modèles de langage et le traitement du langage naturel (NLP). Une expérience en développement web serait un atout supplémentaire. Enfin, un intérêt pour la pédagogie numérique et l'amélioration des méthodes d'enseignement en ligne est fortement souhaité.

5 Informations Pratiques

Ce stage se déroulera sur une durée de 6 mois au sein du laboratoire MOCAH - LIP6 de Sorbonne Université. L'encadrement sera assuré par l'équipe MOCAH, et une gratification sera versée selon la législation en vigueur.

Références

- [1] Chamoun, E., Schlichtkrull, M., & Vlachos, A. (2024). Automated focused feedback generation for scientific writing assistance. ACL 2024.
- [2] Cooper, M., & Klymkowsky, M. (2024). Let us not squander the affordances of LLMs for the sake of expedience. Journal of Chemical Education, 101.
- [3] Jacobs, S., & Jaschke, S. (2024). Leveraging lecture content for improved feedback. CSEE&T 2024.
- [4] Jia, Q., et al. (2022). Automated feedback generation for student project reports : A data-driven approach. Journal of Educational Data Mining, 14(3).
- [5] Reganti, A. N. (2024). Most impactful RAG papers.
- [6] Wongvorachan, T., & Bulut, O. (2022). Feedback generation through artificial intelligence. OTESSA Conference.
- [7] Zerhoudi, S., & Granitzer, M. (2024). PersonaRAG : Enhancing retrieval-augmented generation systems.

**Contacts : Badmavasan Kirouchenassamy (Badmavasan.Kirouchenassamy@lip6.fr)
et Julien Perez (julien.perez@epita.fr)**