

A Hybrid LLM-Driven Recommendation System for Selecting Flutter Packages Using Structured Metadata and Stack Overflow Knowledge

Bandara R.M.I.M.^{1*}, Chathumini K.G.L.¹, and Kumara B.T.G.S.²

¹Department of Computing and Information Systems, Faculty of Computing,
Sabaragamuwa University of Sri Lanka, Sri Lanka

²Department of Data Science, Faculty of Computing,
Sabaragamuwa University of Sri Lanka, Sri Lanka

*rmimbandara@std.appsc.sab.ac.lk

The rapid growth of the Flutter ecosystem has built significant challenges for the developers in choosing the right third-party packages. Existing popularity-based methods fail to capture developer intent or semantic alignment with real-world requirements. To address this gap, this research presents a hybrid LLM-based recommendation system that integrates structured metadata from Flutter repositories with unstructured knowledge from Stack Overflow. With three LLMs comparatively evaluated, GPT-4-Turbo was chosen based on the accuracy, consistency, and low hallucination rate. The final recommendation model uses a weighted scoring method that integrates popularity, recency, likes, downloads, accepted answers and LLM-based semantic relevance, which generates ranked package recommendations based on the user's requirements. Experimental evaluation across multiple test scenarios achieved a precision of 93.2%, recall of 91.5%, F1-score of 92.3%, and an average semantic relevance score of 94/100. Expert evaluation by Flutter practitioners with six to ten years of experience confirmed that 92% of the recommendations aligned with realworld development requirements. Even though the paper focuses on state management packages, the approach methodology can be generalized to other Flutter package categories. The results highlight the potential of LLMbased hybrid systems to improve developer decision-making.

Keywords: *Flutter, Flutter Packages, Large Language Models, Recommendation System, Software Engineering*