

Designing an Effective Architecture for Exploiting Generative AI to Gather Insights and Predictions from Business Data

T.H.E.K.G. Sarathchandra^{1*}, S.R. Ponnampuruma², A.S. Liyanagunawardana³,
S.A.B. Markus⁴

¹*Emojot (Pvt) Ltd, Colombo 3, 00300, Sri Lanka*

²*Emojot (Pvt) Ltd, Colombo 3, 00300, Sri Lanka*

³*Emojot (Pvt) Ltd, Colombo 3, 00300, Sri Lanka*

⁴*Emojot (Pvt) Ltd, Colombo 3, 00300, Sri Lanka*

**kashnikags@gmail.com*

From text to images, music, and beyond, Generative AI is a type of artificial intelligence that not just analyzes existing content but generates content in response to an input. The power of generative AI is vast, particularly across various aspects of business operations. For example, by leveraging businesses's past data, generative AI can analyze customer satisfaction, predict the business's future, and monitor client account health which could be strategically utilized in their favor. One of the critical concerns with this idea is the constraints encountered when feeding large volumes of data to managed AI cloud services. This is a significant concern, especially when deploying generative AI solutions that require extensive datasets to function optimally. Privacy concerns about feeding customer data directly to AI cloud services are also a hot topic when designing this type of tool. To overcome these issues a custom-designed Transformation Layer architecture was implemented. The goal of this architecture was to feed a more optimized version of data to Large Language Models (LLM) without losing its value while respecting the inherent limitations of models and privacy policies. Transformation Layers involve the strategic application of multiple transformations to the raw dataset such as cleanser, anonymizer, text-to-numeric transformer, summarizer, analyzer, etc. By leveraging this approach we were able to increase the volume of data drastically from 100 up to 1500 data records. It also enabled more effective use of generative AI without overwhelming the model to get the most out of data. An added benefit of this approach was the cost-effectiveness when leveraging Generative AI APIs due to the reduced token usage. This Transformation layers architecture will help businesses use generative AI more effectively to turn their business data into actionable insights and help drive customer satisfaction, business growth, and client retention to new heights.

Keywords: Generative AI, Large Language Models, Transformation Layers, Insights, Predictions

Acknowledgement: we express our sincere gratitude to Emojot (Pvt) Ltd for providing resources and facilities.