International Journal of Artificial Intelligence, Data Science, and Machine Learning
Grace Horizon Publication | Volume 6, Issue 1, 115 -123, 2025
ISSN: 3050-9262 | https://doi.org/10.63282/3050-9262.IJAIDSML-V6I1P113

Original Article

Prompt Engineering for LLMs: Real-World Applications in Banking and Ecommerce

Balkishan Arugula
Sr. Technical Architect/ Technical Manager at MobiquityInc(Hexaware), USA.

Received On: 30/11/2024 Revised On: 09/12/2024 Accepted On: 29/12/2024 Published On: 17/01/2025

Abstract: *Big Language Models (LLMs) like GPT have revolutionized computer understanding & creation of human language, therefore generating fresh possibilities in AI-driven problem-solving. As these models develop, timely engineering formulating precise & efficient input prompts has become more important in realizing their full potential. Rapid engineering allows users to guide current models to perform domain-specific actions which are improved by their accuracy and relevance, therefore enabling rather than ground-up development of the latest models. In industries like banking & e-commerce, where complex decision-making processes & huge volumes of unstructured information call for intelligent automation, this approach is showing notable change. Quick engineering helps banks to produce automated reports, identify more intelligent fraud & provide better customer support using natural language interfaces. Prompts are being used by e-commerce systems to improve their product recommendations, personalize shopping experiences, and maximize supply chain communications. These improvements are not merely technical ones; they also change operational effectiveness & also customer experiences. Many actual world case studies show the effectiveness of this approach: for instance, an online retailer improved conversion rates by AI-driven content personalization enabled by careful prompt design, while a financial institution used LLM prompts to greatly lower customer query resolution times. A key difference will be the ability to proactively affect LLM results by quick engineering as companies progressively use artificial intelligence technology. This abstract highlights the need of understanding the nuances of fast design to link broad artificial intelligence capabilities with tailored commercial solutions, thereby transforming complex language models from not only powerful but also quite useful.*

Keywords: *Prompt Engineering, Large Language Models (LLMs), Generative AI, Natural Language Processing, Banking Technology, E-commerce Innovation, AI Chatbots, Fraud Detection, Personalized Recommendations, Credit Risk Analysis, Customer Experience, Regulatory Compliance, Semantic Search, Conversational AI, AI Ethics.*

1. Introduction

Mostly driven by generative models and large language models (LLMs), AI has seen a major progress recently. Human-machine interaction has been transformed by technologies ranging from OpenAI's GPT to Anthropic's Claude to Meta's LLaMA. These models can understand and respond to complex, conversational language, frequently producing outputs that look rather human-like; we are not limited by rigid rules or closely defined inputs. These LLMs are actively revolutionizing companies like banking & e-commerce, two sectors where the need for intelligent automation & customizing has reached hitherto unheard-of degrees; they are not limited to research labs or technology blogs.

This change did not happen suddenly. From the initial phases of natural language processing, when computers followed strict guidelines or relied mostly on their human created traits, we have advanced greatly. Though often useful, these traditional approaches ran up challenges with irony, ambiguity & the always shifting dynamics of human language. Driven by huge datasets and sophisticated training approaches, the arrival of LLMs marked a change from rule-based systems to models that learn from examples. Along with that change came the latest challenge: how best can we interact with these models? Prompt engineering plays this function.

Prompt engineering essentially consists in the creation of inputs that guide a large language model to produce useful, exact, and contextually too many relevant outputs. Though it might appear simple just a matter of asking the relevant questions this is really a dynamic and complicated area. Zero-shot prompting which gives no examples few-shot prompting which offers a limited number of examples to lead the model and chain-of- thought prompting which fosters intermediate thinking processes are among the many techniques employed. Every one of these techniques shapes the way the model understands the nature of the response and the work involved.

What relevance does this have? There are restrictions even on the most powerful LLMs. In the human environment, they do not "comprehend"; in the absence of well constructed prompts, their outputs may be vague, more erroneous, or mismatched with organizational needs. In sectors like banking or e-commerce, this might have major effects deceptive financial analysis, poor customer service interactions, or security flaws resulting from misread information. Between human intent & machine output, prompt engineering acts as a bridge ensuring that the responses of the model are not only technically accurate but also contextually appropriate to the user's goals.

The use of rapid engineering in useful settings especially in banking and e-commerce is investigated in this paper. These sectors were chosen as they represent two extremes of a broad spectrum: one is dynamically customer-oriented while the other is highly regulated & risk-averse. We will look at how financial companies use prompt-based big language models for customer support, regulatory compliance & fraud detection. Within the e-commerce space, we will look at how companies are tailoring product recommendations, answering customer questions, and improving operations all enabled by smart prompts.

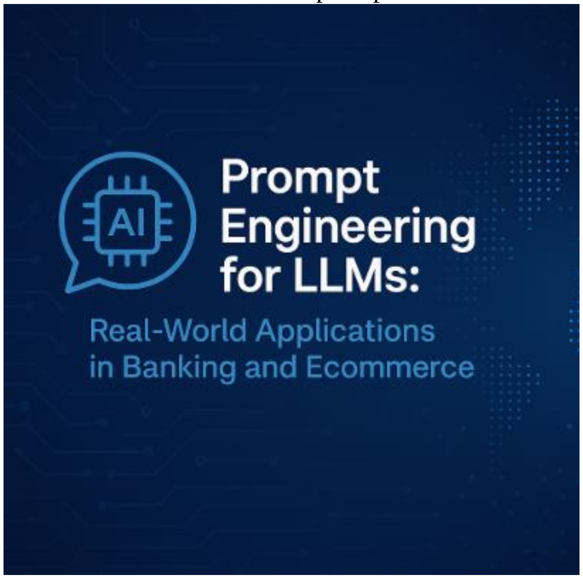

Figure 1: Smart Prompts

We will look at the specific benefits rapid engineering provides more efficiency & better user experiences among many others. We will frankly discuss the difficulties prompt brittleness, model hallucinations, and the requirement of constant iteration. The aim is to show readers options by means of case studies & pragmatic insights, therefore helping them to understand how to approach LLM integration with care and strategy. This article is meant for product managers looking at Chabot applications, IT executives including artificial intelligence into their companies, and those curious about how these latest technologies may affect different industries. Let's look at how the simple act of asking the relevant inquiry might help to enable a domain of intelligent automation and business transformation.

2. Understanding Prompt Engineering

The field of prompt engineering involves maximizing their interactions with large language models (LLMs) like GPT. It is not merely a question of asking the model questions and expecting positive answers; rather, it is a matter of developing those questions to guide the model toward exact, useful & contextually relevant answers. Quick engineering may drastically change the scene in fields like banking & e-commerce, where accuracy and efficiency rule.

2.1. Components of a Workable Prompt

Three basic ideas define excellent fast engineering: structure, clarity & also restrictions.

Clarity means that the instruction is understandable and removes any chance of misinterpretation. Large Language Models perform best under clear, exact instructions. "Summary of this product review," for example, is more successful than "What is your opinion on this?" Structure helps the model to follow a rational development. Consult templates, numbered lists, or bullet points. For instance, asking the model to provide the outcome in a certain tone such as formal or friendly or as a JSON object gives necessary structure.

Constraints serve as model guiding values. One may limit word count, style, or the intended kind of response. "Respond in fewer than 100 words in a professional tone" commands a clear limit that improves the output quality of the model. One of the main determinants of designing prompts is whether to offer simple instructions or examples. For basic or routine chores like writing a greeting, directions can be more sufficient. For more complex or specialized jobs, though such as creating compliance language for banking application illustrative examples help. Presenting the model with many correct outputs raises the possibility of attaining the intended outcomes.

2.2. Approaches for Encouragement

There are many approaches to interact with LLMs depending on your goals & the complexity of the task.

- Zero-shot prompting gives the model only instructions, devoid of any instances, therefore rendering her work. For regular chores like translating or summarizing, this is quick & also successful.
- Few-shot prompting is giving the model several instances to copy. This helps especially in more complex or sector-specific projects. For example, a range of well selected examples may help a model in an e-commerce setting answer customer support queries.
- Fine-tuning is the method of training the model on a particular dataset such that it satisfies your exact requirements. Though it uses more resources, this method offers the best accuracy. Using previous information, banks may hone a model to provide internal compliance reports or spot trends in fraud.
- Apart from these strategies, you may improve the behavior of the model by changing variables such as temperature and maximum tokens.
- Temperature controls chaos. Reduced values such as 0.2 offer responses more focused and predictable ideal for uses where accuracy is too critical, including risk assessment in banking. Higher values like 0.8 may inspire innovation, ideal for e-commerce's marketing or product descriptions.
- The response length is limited by the maximum token count. Particularly important in chatbot or mobile application interfaces, establishing a token limit guarantees respect to character limits or API budgets.

Furthermore affecting the language creation of the model are other elements such as top-p and frequency penalty. Understanding this helps you to more successfully match output to your own needs.

2.3. Frameworks and Instruments

A growing technological environment is developing that simplifies quick engineering & helps LLMs to be included into useful processes.

- Lang Chain shines in combining prompts, memory, and outside tools like databases. Imagine a loan application system employing a big language model customized by their explanations based on client history retrieved from a database. Lang Chain makes the possibility possible.
- Prompt Layer helps track, monitor, and debug prompts. Prompt Layer provides insight into the performance and change of prompts over time for teams working in banking or e-commerce industries where audits & compliance are vital.
- Designed as an interactive tool for playing with prompts and seeing how different variables affect the outcome, the OpenAI Playground For quick testing tone, style, or format prior to use in production, it is ideal.

Often embedded into no-code tools like Zapier or Air table, these solutions allow APIs to be connected. This allows non-technical teams such as an analyst in banking or a marketing manager in e-commerce to utilize LLMs free from coding needs.

2.4. Evaluation Measures

Once one begins using LLMs, one will wish to find their performance efficiency. In this sense, evaluation tools are essential. These help assess your models' reliability & the quality of performance.

- Accuracy gauges whether the model offers correct information. In a financial environment, this may mean precisely spotting fraudulent activity. In e-commerce, it might mean matching a customer inquiry's relevant product category.
- Coherence evaluates the model's logical development of output. Though technically accurate, an inconsistent or conflicting response might nonetheless be confusing.
- Relevance ensures that the answer either satisfactorily answers the question or fixes the problem. In customer service, this is especially important as tailored answers to specific problems are valued over broad recommendations.

The rate of hallucinations shows how often the model creates fault information. In terms of money, illusions might have major effects. Thus, regular assessment and quick development are absolutely required.

3. Applications in Banking

Large language Models (LLMs) are revolutionizing banking processes & improving the efficiency, intelligence, and consumer friendliness of daily chores. Key to this transformation is prompt engineering that is, the ability to create inputs that produce desired outputs from these models. Effective prompts help banks to fully use LLM strengths in many other different fields. We will look at some actual world banking cases where timely engineering is now making an effect.

3.1. Consumer Support Automation

Customer service is one area where LLMs find great use in banking. Conventional Chabot's with their set, rigid responses can irritate users. By means of thorough quick engineering & LLMs, banks might create advanced virtual assistants able to participate in smooth, human-like interactions. The assistant may clearly & amiably answer questions like "What is the procedure to reset my debit card PIN?" or "How do I apply for a home loan?" by creating clues that guide the LLM to understand client intent. These LLMs could handle not just straightforward questions but also more complicated issues requiring context & also memory.

Sentiment-aware dialogues are very effective. By use of suitable cues, the system might detect emotional signals such as perplexity or irritation and respond accordingly. If a user complains about a delayed transfer, for instance, an appropriately motivated LLM may show empathy and suggest bringing the problem up with a human representative. Prompts, like "If user sentiment is negative & the issue pertains to fraud, escalate immediately" within the conversation flows, might drive this escalation logic.

3.2. Detecting and Notifying Fraud

Data analytics has always been the cornerstone of fraud detection; nevertheless, LLMs might provide a different perspective, especially in the interpretation & summary of actual time trends. Prompt engineering lets LLMs examine transaction information and provide understandable descriptions of questionable behavior. For example, a large language model prompted with "Summary unusual activity in this account for the previous 24 hours" can instantly provide insights such, "Multiple high-value withdrawals occurred from overseas IP addresses between 2 a.m. and 4 a.m." instead of a fraud analyst carefully reviewing many other transaction logs.

By examining behavior patterns, large language models may help find more anomalies. By use of questions like "Are there any anomalies in these transactions considering the user's historical behavior?" the model may subjectively analyze information and find disparities that might not be obvious from just numerical analysis. By acting as an intelligent layer for early warnings or review annotations, the LLM improves rather than replaces traditional fraud systems.

3.3. Risk Evaluation and Credit Assessments

Creditworthiness calls for not just numerical facts but also a mix of structured financial information & unstructured components like personal references, social behavior & job histories. Large Language Models might help to examine this thorough review depending on the prompts created to highlight risk aspects & score criteria. Banks provide credit risk summaries combining quantitative & qualitative information using prompts. For instance, one may generate an LLM with "Produce a credit risk assessment for a mortgage loan application," depending on a customer's financial background, job records, and recent spending habits. The output might have sections on risk indicators, debt-to-income ratio, income stability & employment volatility.

One other use is in the interpretation of qualitative information. A loan applicant writes a cover letter mentioning a previous delinquency. By means of judgments anchored on language & logical analysis, a prompt like "Summary the applicant's explanation and evaluate its credibility based on their financial documents" helps the LLM to support human decision-making.

3.4. Regulatory Support & Compliance

Banking compliance covers a wide & complex field. Compliance officials have to understand long-standing legal documents such Anti-Money Laundering (AML) laws and Basel III regulations and match them with daily operations. By use of rapid engineering, Large Language Models may greatly shorten the time needed to negotiate this complexity.

Using prompts that let staff members ask natural language questions such, "What does Basel III stipulate regarding Tier 1 capital?" banks are creating Q&A systems Under AML rules, are we obliged to disclose a $9, 999 transaction? LLMs can answer such questions precisely and clearly with well-crafted prompts & access to accurate regulatory information. Automation of internal audits is another area seeing development. Auditors may order LLMs with instructions such, "Verify compliance of employee onboarding processes with KYC policies," therefore allowing the model to examine their records, spot flaws, and suggest changes. These solutions help companies to stay compliant on their own, thus less depending on legal personnel.

3.5. Internal Knowledge Control

Financial institutions are entities with information-intensity character. Workers typically spend a lot of time looking for the material they need from policy documents to process guidelines. Appropriately asked, LLMs might be powerful internal search engines & content providers. Automation of often asked questions about HR, IT, or banking policies has a pragmatic use. Workers could ask the system questions such as, "How many days of leave may I defer?" or "What is the procedure for reporting insider trading issues?" instead of relying on a specific document. An efficient prompt ensures that the LLM follows the set terminology, picks accurate content, and delivers it correctly.

Data extraction from structured documents as spreadsheets, policy tables, or risk matrices has great use. By use of questions like "Summary of the key differences in mortgage policies, the LLM may extract and evaluate data across various dimensions. Such alarms reduce the need for staff members to manually handle problems & access various files.

4. Applications in E-commerce

Prompt engineering has become a major tool for e-commerce as it helps companies to fully utilize large language models. Quick design is turning AI into a powerful instrument for customer involvement by offering customized shopping experiences and allowing virtual shopping assistants. Let's investigate how it shows up in the world.

4.1 Tailored Product Recommendations

4.1.1 Enhancement for Customer Intent Identification

One very exciting area where rapid engineering shines is in identifying customer needs. Although conventional recommendation systems generally rely on their browsing history or buying patterns, they might ignore the complex details of intent especially when a customer's behavior is inconsistent or exploratory. Intelligent prompts let large language models be improved in order to separate natural language inputs from their intended meaning. For example, a well-crafted prompt may help the model identify the age demographic, relationship context & area of interest to provide tailored suggestions if a user enters, "I seek a gift for my adolescent nephew who is passionate about technology." The model understands the background, which is related to careful rapid engineering, not just aligning keywords.

4.1.2 Linguistically Grounded Suggestion Systems

Next-generation conversational recommendation systems find their cognitive basis in large language models. Think of it as a proactive shopping assistant that probes "Are you seeking something casual or more formal?" Alternatively "Do you have a set budget?" These engines not only provide products but also learn and improve their Recommendations in actual time using well-organized cues. This degree of personalization produces higher customer delight, more conversions, and a more engaging purchase experience.

4.2 Modern Filtering and Search

4.2.1 Asking questions Natural Language Catalogues

Conventional search bars cause problems for clients very regularly. Should they deviate from the exact language expected by the system, they either obtain meaningless results or, at least none at all. LLM prompt engineering helps users accurately express their ideas, hence facilitating natural language search. One may enter, for instance, "red sneakers priced under $100 suitable for running in inclement weather." A good prompt helps the model understand properties (color, price, purpose, and weather suitability) & translate them into ordered searches for system processing. What result follows? Improved search precision, reduced frustration & a flawless checkout experience.

4.2.2. LLM-Enhanced Search Bar Prompts

Search bars may now serve purposes other than just word matching. Models created with prompts correct typographical mistakes provide predictive search suggestions & ask clarifying questions. Should a consumer search "dresses for outdoor wedding," the algorithm may ask: "What is the season or location?" We can help you find solutions fit for the climate. Clever prompt design that lets the LLM predict user needs helps this proactive, interactive search experience to be possible.

4.3. Client Support and Return Management

4.3.1. Dynamic Prompts for Policies on Return/Refunds

In e-commerce, handling returns or refunds sometimes proves difficult. By means of rapid engineering, companies are improving the flow of consumer service contacts. When a customer says, "I received the incorrect item," for example, "May I return it?" the system shouldn't have to negotiate a maze of support documents. A prompt may set the LLM to get the relevant policy, translate it into plain English & provide a direct action like beginning a refund or linking to a live agent. This not only saves customer time but also lessens the load on support staff.

4.3.2. Trigger for Escalation Based on Emotion

Sometimes the way people express themselves matters more than the substance of their words. Large language models designed using prompt-engineering may identify emotional cues. The model may begin an escalation route when a message shows annoyance or urgency, like, "This is the third occurrence, I am quite disappointed." A prompt may ask the model to evaluate more sentiment, context, and previous interactions to decide if the issue should be given top priority or escalated to a human agent. It's AI combined with empathy.

4.4. Product Content and Description Creation

4.4.1. Automated, brand-consistent product copywriting

Creating unique product descriptions in great numbers is an ongoing work. But with appropriately written prompts and LLMs, companies might automate this process while keeping brand consistency in tone and style. One cue may guide the model to generate material in a "luxurious tone," "playful and eccentric voice," or "SEO-optimized format," as needed. It ensures that every product page is painstakingly created instead of being furnished with generic material. For content teams, this saves time; it also enhances the whole shopping experience.

4.4.2. Generation of Interactive Multilingual Content with Prompts

While translating product descriptions is costly & work intensive, global growth requires multilingual support. Often maintaining cultural subtleties, prompt engineering lets LLMs create product content right in the target language right away. Prompts could guide the model to adjust language, idioms & measurements (such as translating inches to centimeters) depending on regional preferences rather than just translating verbatim. It makes the brand basically local regardless of its market.

4.3. Virtual Assistances and Conversational Commerce

4.3.1. Started Discussed Commerce

Conversational buying that is, direct consumer purchases made using chat interfaces is a clear trend. These systems handle payment processes, answer questions, improve upselling & help product discovery by using LLMs trained with specific prompts.

- Imagine this series on WhatsApp:
- Customer: "Desperate for a vegan leather tote bag under $200."
- Bot: “Surely!" Would you go for a sloppy or more disciplined approach?
- Customer: "Casual, probably featuring gold embellishments."
- Bot: "Here are three related options." Would you like to add anything to your basket or see reviews?

Every interaction is guided by invisible signals keeping the discourse more relevant and useful. It comes out as instinctive, reactive, and unique.

4.3.2. WhatsApp, Messenger, and In-App Bots Shopping Assistants

These chat assistants are becoming really important parts of customer engagement; they are not merely useful tools. Over all talk, prompt-engineered LLMs may provide size recommendations, gift ideas, abandoned cart warnings & checkout assistance. Businesses might make these bots less script-like and more like actual shopping companions by designing signals that direct the conversation and allow for flexibility.

5. Case Studies

5.1. A Leading Bank’s AI Chabot

Managing huge scale consumer service operations has always been difficult for financial organizations. One well-known bank received too many customer calls, hugely related to routine issues. Even with a traditional chatbot present, the containment rate that is, the percentage of questions answered entirely without human intervention remains poor. Longer wait times, more staff costs & more customer dissatisfaction followed from this.

To address this, the bank used a finely calibrated Large Language Model (LLM) applied with a few-shot prompting approach. Rather than retraining the model from the ground up, they gave the LLM a small collection of carefully selected cases showing the proper handling of regular customer questions. These little instructions within the prompt helped the model to grasp tone, structure & expected outcomes for every kind of query.

The change was really too quick. The chatbot began handling a wider range of questions with much more refined accuracy & grace. Often foreseeing more questions, the bot skillfully and assertively answered customer queries about credit card eligibility, transaction delays & also PIN resets. Along with lowering running expenses, this reduced the amount of calls needing escalation to human agents. Levels of customer satisfaction also clearly rose. The organic involvement and quick fixes prized by clients let the support personnel focus on more complex or sensitive issues. Especially in highly regulated sectors like banking, this research underlined the effectiveness of few-shot prompting as a scalable, affordable method for improving their legacy support systems.

5.2. E-commerce platform

The shopping experience might be much influenced by the efficiency of search features. One recurring problem an e-commerce platform found was a significant number of customers stopping their search sessions without making a purchase. Even with a huge range of products, consumers often battled to find their preferred ones, sometimes because of conflicting or inconsistent search results.

The company rebuilt their search engine using llMs. Instead of relying only on keyword matching, they used semantic search enabled by rapid integration with a language model. The questions were meant to help a user determine the reason for their search & provide matching answers. When a user searches "shoes for standing all day," for instance, traditional search engines could provide an odd mix of sandals and sneakers. Emphasizing orthopedic shoes, work boots with additional padding, and highly rated walking shoes, the latest LLM-supported engine identified the intent comfort during lengthy standing.

This minor change in the way the results were curated had a big effect. Conversion rates rose 18% shortly after the program started. Customers were buying, frequently with less conflict, not merely browsing. It went beyond simple improvement of search output. The questions guided the model to provide product descriptions, highlight relevant user comments & suggest like products. This produced a useful & individualized shopping assistant experience. A remedy for poor search performance turned into a competitive advantage that improves the digital experience to match the friendly, tailored character of in-store encounters.

5.3. Generation AI-Based Fraud Surveillance

One important and ever changing endeavor is fraud detection. Financial institutions are still alert; nonetheless, the volume of transactions makes human judgment impossible. For one Fintech Company, the limitation was clear: analysts were spending a lot of time, usually under great time pressure, closely examining transaction histories & customer profiles to identify their patterns. To improve both speed & accuracy, the company used an LLM-powered solution with chain-of-thought prompting. This approach drives the model to meticulously go over its thinking process, just as a human might examine their questionable behavior.

A transaction started from an odd location not long after a card was used overseas. The model would provide a rational justification instead of have sine the transaction: "This acquisition from London occurs two hours after a charge in New York." Given the time zone difference and lack of past travel logs, this is most certainly extraordinary. An analyst might then quickly review this summary, usually needing just a cursory check to confirm the later activities. What is the result? Managing fraud now takes less than half of what it did fifty years ago. The method allowed teams without losing completeness quick replies by combining complex patterns into brief summaries.

The language model also helped to spot "gray area" situations transactions with slight variances but minimal relevance. These had gone unseen before. The clear thought paths of the model helped them to be presented more regularly for human assessment. This example showed how fast engineering may help rather than replace analysts by giving them a more acute & quick view for their work.

6. Future Trends and Outlook

We are clearly moving toward more autonomous & intelligent systems as fast engineering develops. One of the most exciting developments is the rise of autonomous agents systems able to complete complex, multi-stage chores with little human involvement. These agents think, adapt & run independently utilizing more dynamic inputs and memory, so they require much advanced quick engineering. In sectors such as banking & e-commerce, this might provide AI systems that independently monitor consumer care activities, spot fraud, or enable customized purchase experiences, all in actual time.

One important tendency is the creation of self-enhancing & more adaptive cues. These cues change depending on user reaction & environmental cues to improve their efficacy over time instead of depending only on their fixed instructions. Through constant matching with user needs, this feedback loop helps systems to improve their accuracy & also efficiency. Imagine a recommendation engine that improves its decisions depending on complex behavioral clues or a digital banking assistant that fits your financial questions and increases its value with every interaction.

The ability of fast engineering to combine many modalities marks its future. Multimodal cues that which combine text with images, graphs, or music are becoming really more powerful. In e-commerce, this might apply to AI that evaluates customer comments or offers improved by these recommendations based on their knowledge of product descriptions in concert with images. In banking, it might examine dashboards & report information to provide strategic financial insights.

Simultaneously emerging is the regulatory environment. Governments and companies are starting to set standards regarding the design and interpretation of prompts to ensure the ethical and correct use of artificial intelligence. This is crucial especially in highly regulated industries like banking, where responsibility and transparency rule most of all. Rapid engineering is becoming a basic ability, hence its future will be shaped not only by creativity but also by the structures set to guarantee its ethical usage.

7. Conclusion

Examining useful applications of prompt engineering for Large Language Models (LLMs) in banking & e-commerce shows that the careful design of prompts may significantly affect the usefulness, accuracy & the reliability of AI systems. Rapid engineering improves customer service efficiency in banks, speeds fraud detection & supports regulatory compliance. In e-commerce, well crafted prompts are improving product recommendations, enabling better customizing, and increasing customer contact across systems. The unifying thread running throughout these successes is the need of fast engineering as the means of bridging natural AI capacity with practical economic advantage. It's not simply about letting LLMs generate meaningful language; it's also about matching AI behavior to context, goals & also user expectations. A little change in language or structure may greatly affect the output of the model, therefore transforming a generic tool into a strategic resource.

This suggests the growing demand for customized fast engineering models within several sectors. Unlike e-commerce, which thrives on creativity and flexibility, banking is defined by strict rules and a demand for precision and requires different signals. One general strategy is insufficient. Companies should make investments in developing best practices particular to their industry and customer needs as well as quick libraries. As operational artificial intelligence develops, the requirement of proper application becomes first priority. Prompt engineering is also important as it helps models to avoid biased, harmful, or misleading responses via cautious supervision. This is a shared responsibility among engineers, domain experts, and ethical leaders all across teams. Future success of artificial intelligence depends on our interactions with the models as much as on their performance. The language of our modern day is prompt engineering, so businesses must interact with it deliberately, ethically, and fluently.

References

[1] Nananukul, Navapat, Khanin Sisaengsuwanchai, and Mayank Kejriwal. "Cost-efficient prompt engineering for unsupervised entity resolution in the product matching domain." *Discover Artificial Intelligence* 4.1 (2024): 56.

[2] Fan, Minghong. "LLMs in Banking: Applications, Challenges, and Approaches." *Proceedings of the*

International Conference on Digital Economy, Blockchain and Artificial Intelligence. 2024.

[3] Koul, Nimrita. *Prompt Engineering for Large Language Models*. Nimrita Koul, 2023.

[4] Yasodhara Varma. "Modernizing Data Infrastructure: Migrating Hadoop Workloads to AWS for Scalability and Performance". *Newark Journal of Human-Centric AI and Robotics Interaction*, vol. 4, May 2024, pp. 123-45

[5] Zhao, Hongke, et al. "A comprehensive survey of large language models in management: Applications, challenges, and opportunities." *Challenges, and Opportunities (August 14, 2024)* (2024).

[6] Atluri, Anusha, and Vijay Reddy. "Total Rewards Transformation: Exploring Oracle HCM's Next-Level Compensation Modules". *International Journal of Emerging Research in Engineering and Technology*, vol. 4, no. 1, Mar. 2023, pp. 45-53

[7] Cheung, Ming. "A Reality check of the benefits of LLM in business." *arXiv preprint arXiv:2406.10249* (2024).

[8] Castelnovo, Alessandro, et al. "Augmenting XAI with LLMs: A Case Study in Banking Marketing Recommendation." *World Conference on Explainable Artificial Intelligence*. Cham: Springer Nature Switzerland, 2024.

[9] Sangeeta Anand. "Fully Autonomous AI-Driven ETL Pipelines for Continuous Medicaid Data Processing". *JOURNAL OF RECENT TRENDS IN COMPUTER SCIENCE AND ENGINEERING (JRTCSE)*, vol. 13, no. 1, Feb. 2025, pp. 108–126

[10] Khan, Ian. *The quick guide to prompt engineering: Generative AI tips and tricks for ChatGPT, Bard, Dall-E, and Midjourney*. John Wiley & Sons, 2024.

[11] Yasodhara Varma. "Managing Data Security & Compliance in Migrating from Hadoop to AWS". *American Journal of Autonomous Systems and Robotics Engineering*, vol. 4, Sept. 2024, pp. 100-19

[12] Sangeeta Anand, and Sumeet Sharma. "Scalability of Snowflake Data Warehousing in Multi-State Medicaid Data Processing". *JOURNAL OF RECENT TRENDS IN COMPUTER SCIENCE AND ENGINEERING (JRTCSE)*, vol. 12, no. 1, May 2024, pp. 67-82

[13] Vega Carrazan, Pablo Federico. *Large Language Models Capabilities for Software Requirements Automation*. Diss. Politecnico di Torino, 2024.

[14] Talakola, Swetha. "Microsoft Power BI Performance Optimization for Finance Applications". *American Journal of Autonomous Systems and Robotics Engineering*, vol. 3, June 2023, pp. 192-14

[15] Paidy, Pavan. "AI-Augmented SAST and DAST Integration in CI CD Pipelines". *Los Angeles Journal of Intelligent Systems and Pattern Recognition*, vol. 2, Feb. 2022, pp. 246-72

[16] Kodete, Chandra Shikhi, et al. "Robust Heart Disease Prediction: A Hybrid Approach to Feature Selection and Model Building." *2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)*. IEEE, 2024.

[17] Mehdi Syed, Ali Asghar, and Shujat Ali. "Kubernetes and AWS Lambda for Serverless Computing: Optimizing Cost and Performance Using Kubernetes in a Hybrid Serverless Model". *International Journal of Emerging Trends in Computer Science and Information Technology*, vol. 5, no. 4, Dec. 2024, pp. 50-60

[18] Johnsen, Maria. *Large language models (LLMs)*. Maria Johnsen, 2024.

[19] Vasanta Kumar Tarra and Arun Kumar Mittapelly. "The Role of Generative AI in Salesforce CRM: Exploring How Tools Like ChatGPT and Einstein GPT Transform Customer Engagement". *JOURNAL OF RECENT TRENDS IN COMPUTER SCIENCE AND ENGINEERING (JRTCSE)*, vol. 12, no. 1, May 2024, pp. 50-66

[20] Veluru, Sai Prasad, and Swetha Talakola. "Continuous Intelligence: Architecting Real-Time AI Systems With Flink and MLOps". *American Journal of Autonomous Systems and Robotics Engineering*, vol. 3, Sept. 2023, pp. 215-42

[21] Atluri, Anusha, and Vijay Reddy. "Cognitive HR Management: How Oracle HCM Is Reinventing Talent Acquisition through AI". *International Journal of Artificial Intelligence, Data Science, and Machine Learning*, vol. 6, no. 1, Jan. 2025, pp. 85-94

[22] Arawjo, Ian, et al. "Chainforge: A visual toolkit for prompt engineering and llm hypothesis testing." *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024.

[23] Veluru, Sai Prasad. "Streaming MLOps: Real-Time Model Deployment and Monitoring With Apache Flink". *Los Angeles Journal of Intelligent Systems and Pattern Recognition*, vol. 2, July 2022, pp. 223-45

[24] Kupanarapu, Sujith Kumar. "AI-POWERED SMART GRIDS: REVOLUTIONIZING ENERGY EFFICIENCY IN RAILROAD OPERATIONS." *INTERNATIONAL JOURNAL OF COMPUTER ENGINEERING AND TECHNOLOGY (IJCET)* 15.5 (2024): 981-991.

[25] Chaganti, Krishna Chaitanya. "Ethical AI for Cybersecurity: A Framework for Balancing Innovation and Regulation." *Authorea Preprints* (2025).

[26] Sokolovas, Manvydas. *Investigation of process automation with large language models*. Diss. Vilniaus universitetas., 2024.

[27] Sangaraju, Varun Varma. "UI Testing, Mutation Operators, And the DOM in Sensor-Based Applications.

[28] Mehdi Syed, Ali Asghar. "Zero Trust Security in Hybrid Cloud Environments: Implementing and Evaluating Zero Trust Architectures in AWS and On-Premise Data Centers". *International Journal of Emerging Trends in Computer Science and Information Technology*, vol. 5, no. 2, Mar. 2024, pp. 42-52

[29] Alto, Valentina. *Building LLM Powered Applications: Create intelligent apps and agents with large language models*. Packt Publishing Ltd, 2024.

[30] Talakola, Swetha. "Enhancing Financial Decision Making With Data Driven Insights in Microsoft Power BI". *Essex Journal of AI Ethics and Responsible Innovation*, vol. 4, Apr. 2024, pp. 329-3

[31] Chaganti, Krishna Chaitanya. "AI-Powered Patch Management: Reducing Vulnerabilities in Operating Systems." *International Journal of Science And Engineering* 10.3 (2024): 89-97.

[32] Amini, Reza, and Ali Amini. "An overview of artificial intelligence and its application in marketing with focus on large language models." *International Journal of Science and Research Archive* 12.2 (2024): 455-465.

[33] Kumar Tarra, Vasanta, and Arun Kumar Mittapelly. "AI-Driven Lead Scoring in Salesforce: Using Machine Learning Models to Prioritize High-Value Leads and Optimize Conversion Rates". *International Journal of Emerging Trends in Computer Science and Information Technology*, vol. 5, no. 2, June 2024, pp. 63-72

[34] Paidy, Pavan. "Scaling Threat Modeling Effectively in Agile DevSecOps". *American Journal of Data Science and Artificial Intelligence Innovations*, vol. 1, Oct. 2021, pp. 556-77

[35] Irshad, M. "Exploring LLMS, A Systematic Review with SWOT Analysis." *J Artif Intell Mach Learn & Data Sci 2024* 2.4: 1749-1766.

[36] Bustos, Juan Pablo, and Luis Lopez Soria. *Generative AI Application Integration Patterns: Integrate large language models into your applications*. Packt Publishing Ltd, 2024.

[37] Kodi, D. (2024). "Automating Software Engineering Workflows: Integrating Scripting and Coding in the Development Lifecycle ". Journal of Computational Analysis and Applications (JoCAAA), 33(4), 635–652.