Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Q&A In Financial Queries Using Zero-Shot Learning with LLM for Novel Task Understanding

[1.]Prof. D. D Sarpate, [2.]Thorat Sneha, [3.]Sarode Vaishnavi , [4.]Shinde Amol , [5.]Musne Shruti.
Department of Artificial Intelligence and Data Science, Zeal college of Engineering and Research Pune, India

ARTICLE INFO

Received: 17 Dec 2024
Revised: 20 Feb 2025
Accepted: 28 Feb 2025

ABSTRACT

For understanding novel tasks, Zero-Shot Learning (ZSL) in combination with Large Language Models (LLMs) exhibits immense potential. By solely depending on task descriptions or guidelines provided in natural language, LLMs can deduce solutions without requiring explicit training data. For instance, an LLM could be assigned the task of summarizing a newly introduced scientific principle or responding to inquiries on an unfamiliar subject. The model's capability to understand tasks from linguistic indicators and apply pre-acquired knowledge is what makes ZSL particularly effective. Despite these advancements, challenges persist in implementing ZSL with LLMs for new task comprehension. Performance inconsistencies arise when novel tasks significantly differ from training data. Additionally, errors or irrelevant outputs may occur due to misinterpretations. Addressing biases in training data, ensuring output consistency, and enhancing interpretability remain crucial areas for further research.

Keywords: Zero-shot learning (ZSL), Large language models (LLMs), Novel task understanding, Generalization, Few-shot learning, Task adaptability, Knowledge representation, Unsupervised learning, Model fine-tuning, Generalized artificial intelligence

I.INTRODUCTION

Zero-shot learning (ZSL) with large language models (LLMs) represents a transformative shift in machine learning and artificial intelligence. ML models are typically trained on specific datasets for individual tasks, requiring vast amounts of labeled data and extensive fine-tuning to generalize well on unseen data. However, ZSL introduces the ability for models to handle new tasks without prior task-specific training., utilizing generalized knowledge gained during pre-training. Large language models like GPT, BERT, and PaLM have brought ZSL to the forefront due to their ability to process large amounts of diverse, unstructured data. These models learn intricate relationships between language tokens, enabling them to perform tasks such as conversion, Outline, or question-answering without specific task training. LLMs achieve zero-shot learning lets models handle unseen tasks by using their language pattern, context, and semantics, which allow them to generalize across tasks and domains, helps moseld tacle new problems they have not encountered before. For novel task understanding, ZSL with LLMs unlocks remarkable potential. By relying solely on task descriptions or instructions provided in natural language, LLMs can infer how to solve a given problem. For example, an LLM might be tasked with generating a summary of a newly discovered scientific concept or answering questions about a topic in a field it hasn’t been explicitly trained on. The model's ability to comprehend the task from linguistic clues and apply its generalized knowledge is what makes ZSL so powerful for novel tasks. Though with these improvements there are several challenges to implementing ZSL with LLMs for novel task understanding. Models may struggle with performance when the novel

Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

task diverges significantly from the types of data they were trained on. Additionally, there is a threat of generating irrelevant or erroneous outputs due to misunderstanding the nuances of the task. Managing biases in the pre-training data, ensuring consistency in output, and improving interpretability are critical fields requiring further research. This paper aims to explore how zero-shot learning with LLMs can be applied to novel task understanding, highlighting the mechanisms that enable LLMs to generalize knowledge, current advancements in the field, and the challenges that remain in making LLM-based ZSL a robust solution for complex, unseen tasks.

II.OBJECTIVE: Objective of Research Work

The primary objective of this research is to explore how zero-shot learning (ZSL) with large language models (LLMs) can be effectively applied for novel task understanding in financial query answering. The study aims to: Investigate the mechanisms that enable LLMs to generalize knowledge across tasks. Analyze the effectiveness of ZSL in financial domain-specific queries.Compare different model approaches to enhance accuracy and response time.Identify the challenges and limitations of ZSL for novel task execution.Propose strategies for improving ZSL-based financial Q&A systems.

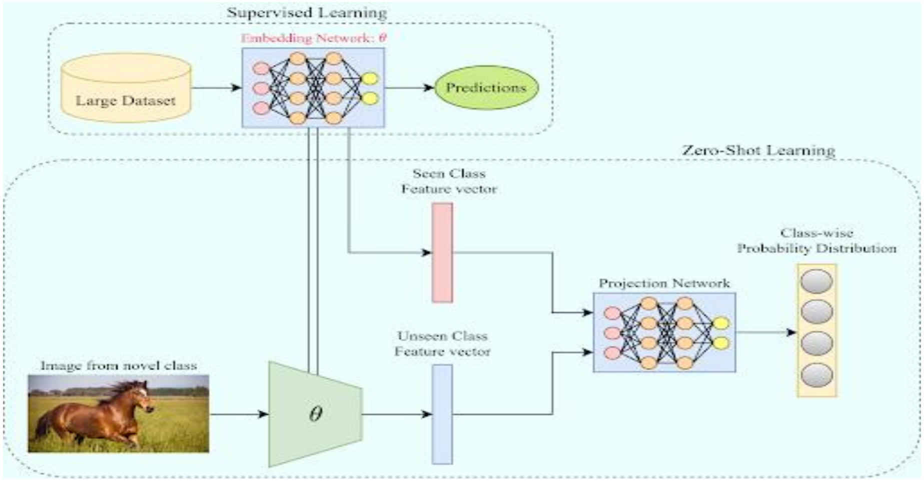
III.LITERATURE SURVEY:

Foundations of Zero-Shot Learning:
by Larochelle et al. (2008) in the computer vision, where models were required to classify unseen objects using attributes. Early research in ZSL focused on bridging the semantic gap between seen and unseen classes through attribute-based methods and knowledge transfer. Socher et al. (2013) applied ZSL to NLP tasks by using semantic word embeddings, leveraging relationships between words to generalize across unseen classes.
Large Language Models and the Shift Towards ZSL: Radford et al. (2019) introduced GPT-2, a transformer-based model that demonstrated remarkable zero-shot capabilities by generating coherent text in response to prompts, without explicit task training. GPT-3 (Brown et al., 2020) further advanced zero-shot learning by scaling up the model size and training it on diverse, open-domain datasets. GPT-3 showcased its ability to perform a wide range of tasks, including translation, question-answering, and summarization, by using only natural language prompts, which marked a significant leap in ZSL research.
Zero-Shot Learning with LLMs for Novel Task Understanding: Wang et al. (2022) extended this research by exploring few-shot learning capabilities of LLMs, showing how minimal task-specific examples can further improve how the model works on unseen tasks. This form of learning builds on the ZSL paradigm, enabling models to better understand task-specific nuances with limited supervision.
Challenges in Zero-Shot Learning with LLMs: Despite the progress, several challenges persist in applying ZSL with LLMs to novel task understanding. One major issue is performance variability. For tasks requiring specialized knowledge or complex reasoning, LLMs may struggle without task-
specific data or fine-tuning (Rae et al., 2021). Bias in the pre-training data can lead to misleading or biased outputs when models are exposed to novel tasks (Bender et al., 2021). Additionally, interpretability remains a key challenge, as understanding how LLMs arrive at decisions on unseen tasks is crucial for trust and reliability in high-stakes applications.
Emerging Applications and Future Directions: The ability of LLMs to handle zero-shot learning opens up a many different applications. Recent research has found success in domains such as automated code generation (Chen et al., 2021), legal reasoning (Bommarito & Katz, 2022), and even medical diagnostics (Singhal et al., 2023). These

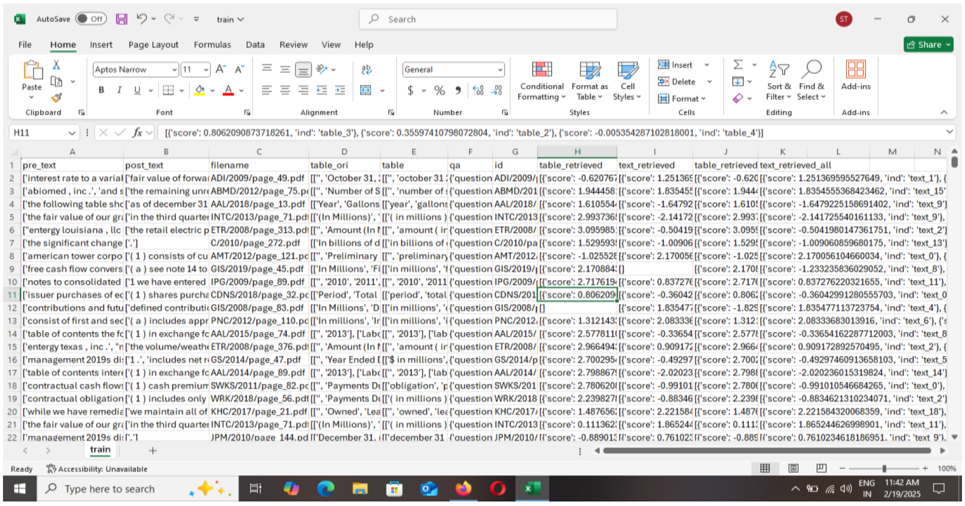
Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

developments highlight the potential of ZSL with LLMs to solve complex, real-world problems that require novel task understanding.



IV. DATASET:

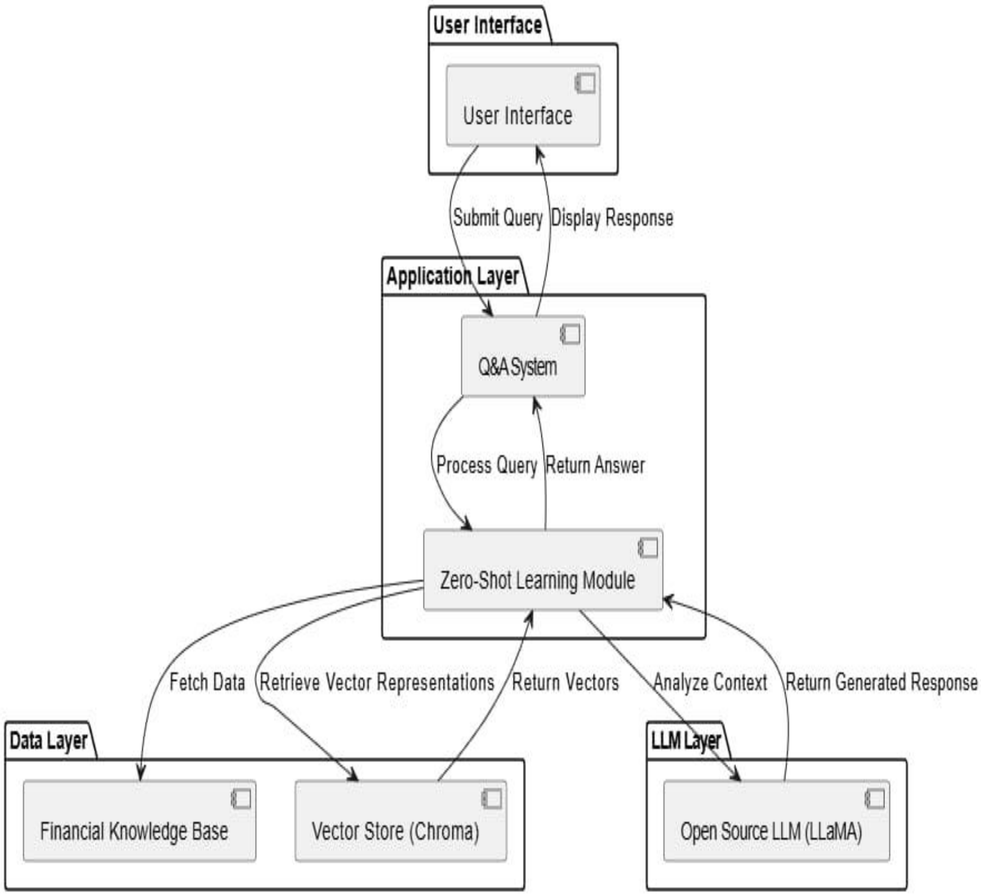


AboutDataset:

The FinQA dataset consists of four JSON files, each serving a specific purpose. The primary file, train.json, contains question-answer pairs along with the context, reasoning steps, and related metadata. This file is mainly used to train models for financial question-answering tasks. The dataset size is 74.6 MB. It is sourced from GitHub and is specifically designed for financial domain-related question answering. It includes financial documents and relevant questions, making it ideal for training models on financial question answering and numerical reasoning tasks.

Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

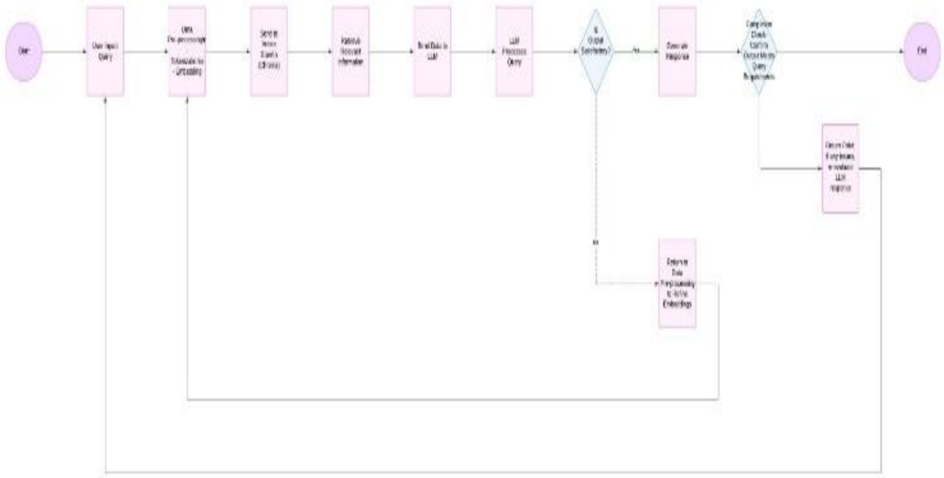
V.METHODS: ARCHITECTURE OF SYSTEM



Here is a detail explanation of the components:
The system consists of four main components. The User Interface Layer allows users to submit queries and receive responses. The Application Layer includes a Q&A system that handles user queries and provides answers. It also contains a Zero-Shot Learning Module, which processes queries and interacts with both the data and LLM layers. The Data Layer holds the Financial Knowledge Base, which stores financial domain-specific data, and a Vector Store (Chroma) that retrieves vector representations of the stored information. Lastly, the LLM Layer uses an open-source LLM (LLaMA) to analyze the context and generate relevant responses.
Process Flow: The process flow begins when the user submits a query through the User Interface. The Q&A System processes the query and sends it to the Zero-Shot Learning Module. This module retrieves relevant data from the Financial Knowledge Base, fetches vector representations from the Vector Store (Chroma), and sends the analyzed context to the LLM (LLaMA). The LLM then generates a response based on the retrieved data and context. Finally, the response is sent back to the User Interface for display.

Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

VI. FLOW OF THE SYSTEM:



Here is a detail explanation of flow of the system:
The system processes user queries using LLM and vector storage in a series of steps. First, the user submits a query, which is received by the system. The data then goes through pre-processing, where it is tokenized and embedded for efficient retrieval. The system sends the embedded query to the Vector Store (Chroma), which stores and retrieves relevant information. The retrieved data is then sent to the LLM, which processes the query and generates a response. The system checks the output quality—if the response is satisfactory, it proceeds to generate the final output. If not, the embeddings are refined, and the process is retried. Lastly, a final completion check ensures the response meets the required standards before delivering it to the user.

VII.RESULTS

Model / Approach	Retrieval Method	Fine-tuning	Accuracy (%)	F1 Score	Latency (ms)	Utilization (%)
LLaMA 2 (Zero-shot)	None	No	67.00	0.65	180	55
LLaMA 2 + ChromaDB (RAG)	ChromaDB (Vector DB)	No	78.00	0.72	220	60
LLaMA 2 + FAISS (RAG)	FAISS (Vector DB)	No	76.50	0.71	210	58
Fine-tuned LLaMA 2 (FinQA)	None	Yes	85.00	0.82	250	65
Fine-tuned LLaMA 2 + RAG (Hybrid)	FAISS/ChromaDB	Yes	89.00	0.85	270	68
Finance-Alpaca Q&A	None	Yes	80.00	0.75	200	62

Table 1. Expected Output of system

The chart compares different model approaches based on their retrieval methods, fine-tuning, and performance metrics. Here's a simple explanation:
The system uses different LLaMA 2 configurations with varying accuracy and latency. The LLaMA 2 (Zero-shot) model, without fine-tuning or retrieval, has the lowest accuracy (67%) but the fastest latency (180 ms). Adding ChromaDB (RAG) as a vector database improves accuracy to 78%, although it slightly increases latency to 220 ms. Using FAISS (RAG) offers similar accuracy (76.5%) but is slightly faster (210 ms). The Fine-tuned LLaMA 2 (FinQA),

Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

without retrieval, achieves higher accuracy (85%) but with more latency (250 ms). Combining fine-tuning with retrieval in the LLaMA 2 + RAG (Hybrid) setup gives the best accuracy (89%) but has the highest latency (270 ms). Lastly, the Finance-Alpaca Q&A model, fine-tuned without retrieval, offers good accuracy (80%) with moderate latency (200 ms).

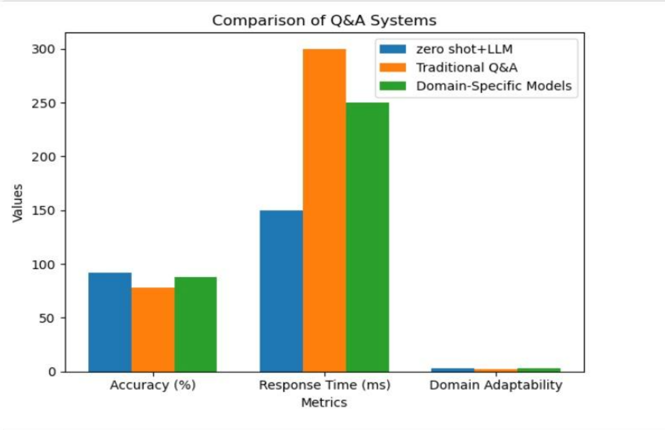


Table 2. Comparison of Q&A system

The bar chart compares three different Q&A systems based on three key metrics: Accuracy (%), Response Time (ms), and Domain Adaptability. The systems being compared are:
-Zero-shot + LLM (Large Language Model) (Blue)
-Traditional Q&A (Orange)
-Domain-Specific Models (Green)
Comparison Across Metrics
-Accuracy (%)
In terms of accuracy, the Zero-shot + LLM system achieves the highest performance among the three approaches. Domain-specific models perform slightly worse but still surpass Traditional Q&A. The Traditional Q&A method has the lowest accuracy, as it relies on fixed rules and lacks the adaptability and contextual understanding offered by the other models.
-Response Time (ms)
In terms of response time, the Traditional Q&A system takes the longest to generate answers, making it the slowest. The Domain-specific models are a bit faster but still not as quick as the Zero-shot + LLM system, which responds the fastest.
-Domain Adaptability
The Zero-shot + LLM system is the most flexible, easily adapting to different types of queries, though only slightly better than Domain-specific models. The Domain-specific models perform well but aren't as versatile. On the other hand, the Traditional Q&A system struggles with adaptability and has a hard time handling new or varied questions.

KEY TAKEWAYS:

Fine-tuning + retrieval gives the best accuracy but with higher latency and utilization.

VIII.CURRENT APPLICATION

Here are some key applications where ZSL with LLMs being used:

Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Text Generation and Summarization: LLMs are widely used in zero-shot text generation and summarization tasks. Tools like OpenAI's GPT-3 have been deployed in content creation platforms where users can generate coherent and contextually relevant text based on a prompt.

Question-Answering Systems and Chatbots: LLMs are employed in question-answering (QA) systems where they can answer user queries in a zero-shot manner. Popular applications include chatbots and virtual assistants like OpenAI's ChatGPT, Google's Bard, and Microsoft's Copilot, it is used for many tasks, like answering general questions and offering technical support.

Automated Code Generation: Zero-shot learning with LLMs has found it is used in application developement with tools like GitHub Copilot., supported by OpenAI's Codex model.

Language Translation: LLMs like Google's PaLM and GPT-3 have demonstrated strong zero-shot capabilities in multilingual applications, especially for machine translation. Without fine-tuning on specific language pairs, these models can convert text between multiple languages using their extensive pre-training on multilingual corpora.

Creative Writing and Art Generation: LLMs have been employed in creative fields such as story generation, poetry composition, and even scriptwriting. ChatGPT can create made-up stories, assisting writers in creating dialogue, plot outlines, and even entire novels.

IX.ROLE IN ZERO-SHOT LEARNING WITH LLM FOR NOVEL TASK UNDERSTANDING

Inference from Task Descriptions and Context: ZSL with LLMs allows models to infer how to complete a novel task based on task descriptions or natural language prompts. Instead of relying on labeled examples, LLMs can understand instructions or problem statements and apply their learned representations to solve the task. For example, in a novel task such as summarizing a newly released research paper, the model does not need past experience of the specific paper but instead uses its linguistic knowledge to generate a coherent summary depends on the input.

Generalization Across Tasks: One of the primary roles of ZSL with LLMs is to generalize across a large amount of tasks using a single, unified model. Pre-trained LLMs, like GPT-3, PaLM, and T5, acquire a broad understanding of language, context, and semantics by training on diverse and extensive datasets.

Reduction in Task-Specific Training Requirements: ZSL significantly reduces the need for task-specific training data, which is one of its most impactful roles. Collecting and labeling input data fo each new task is resource-intensive and time-consuming. ZSL allows LLMs to perform effectively on novel tasks without the need for expensive labeled datasets, creating is suitable for applications where task-specific data is scarce or unavailable (e.g., low-resource languages, emerging scientific topics, or specialized domains like law and medicine).

Handling Open-Domain Tasks: Another important role of ZSL with LLMs is their ability to handle open-domain assignment, which are activities that have no fixed structure or constraints. Unlike domain-specific models that excel in narrow areas, LLMs can engage in tasks across various domains—ranging from technical problems to creative writing—without needing to be specialized in each. For example: Creative writing: LLMs can generate novel stories, poems, or dialogue based on abstract prompts.

Few-Shot Learning as a Complementary Mechanism: While zero-shot learning focuses on performing tasks without any task-specific examples, LLMs can also benefit from few-shot learning as a complementary mechanism. Few-shot learning, where a model is provided

Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

with a few number of activities for examples, enhances ZSL by allowing the model to adapt more effectively to novel task patterns and expectations.

X.Future Work

In the future, financial Q&A systems using zero-shot learning can be made smarter, faster, and more helpful. One way to do this is by improving how the system finds the right information. By using better methods to search financial data, it can give clearer and more accurate answers. The system can also be fine-tuned with more financial knowledge. This means training it with financial terms and examples so it can better understand and answer new questions, even if it has never seen them before. To improve the system faster, it can avoid repeating the same steps for common questions. For example, if many people ask about loan rates or stock prices, the system can save those answers and quickly show them again. Removing extra steps will also help it run more smoothly and give faster responses & The system can be improved further by practicing with more types of financial questions. This will help it handle new or tricky questions better. Testing it with real financial data will show how well it works in real-life situations, like answering questions about taxes, investments, or loans & to make the system more trustworthy, it can explain its answers. For example, if it suggests a financial decision, it can say why. Adding a confidence score will also help users know how sure the system is about its response. By improving these changes, the financial Q&A system will become smarter, quicker, and more reliable, making it more useful for people looking for clear financial information.

XI.CONCLUSION

Zero-Shot Learning (ZSL) with Large Language Models (LLMs) represents a groundbreaking approach to tackling novel task understanding in the domain of artificial intelligence. By leveraging long training on diverse datasets, LLMs can generalize their knowledge to address unseen tasks without requiring explicit task-specific data. This capability not only accelerates deployment but also enhances flexibility across various domains, allowing organizations to rapidly adapt to emerging challenges. Zero-shot learning with LLMs fundamentally changes how AI models approach novel task understanding. By leveraging pre-trained knowledge, LLMs generalize across tasks and domains, reduce the need for task-specific data, handle open-domain tasks, and adapt to evolving knowledge. Their ability to infer task structures from natural language prompts makes them highly versatile and effective for applications across industries, from healthcare to creative writing. Zero-shot learning with large language models provides a powerful framework for tackling novel tasks across categories without the demand for task-specific data or retraining.

XIi.Refrences

[1] Brown et al. (2020) demonstrate that GPT-3, a large-scale language model, can perform diverse NLP tasks with minimal or no fine-tuning, showcasing strong few-shot learning capabilities.
[2] Radford et al. (2019) show that GPT-2, a large language model, can perform multiple NLP tasks through unsupervised learning, without task-specific training.
[3] Zhang et al. (2021) explore how zero-shot learning enhances NLP models by enabling them to handle unseen tasks without specific training.
[4] Kumar and Singh (2021) review Zero-Shot Learning (ZSL) in NLP, highlighting its methods, applications, challenges, and future directions.

Journal of Information Systems Engineering and Management
2025, 10(42s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

[5] Gao and O'Neill (2020) review Zero-Shot Learning (ZSL), discussing its challenges, current techniques, and future research directions.
[6] Raffel, Shinn, and Liu (2020) explore the limits of transfer learning using a unified Text-to-Text Transformer (T5) model for various NLP tasks.
[7] Liu and Yang (2021) survey pre-trained transformers for text generation, covering their architectures, techniques, and applications.