Research Report on GNN, LLM, LSTM and Information Retrieval

1. Introduction: What I Studied

I started my research by learning about GNN, LLM, LSTM that help computers understand and process information better.

- **Large Language Models (LLMs)**: These are smart programs that read and understand a lot of text. They can answer questions, summarize information, and even write stories.
- **Long Short-Term Memory (LSTM)**: This is a type of neural network that helps computers remember things over time, making it useful for speech recognition and time-series predictions.
- **Graph Neural Networks (GNNs)**: These models help computers understand relationships between different pieces of data, like how friends are connected in a social network.
- **Information Retrieval (IR)**: Also learned about IR. This is a technique that helps find the most relevant information when someone searches for something online.

After learning about these, I searched to find a model that combined some of these technologies to create something that show the importance and usefulness of these models.

---

2. The Best Model I Found

Among all the models I studied, one stood out the most: **Multilingual Information Retrieval with Large Language Models**. This model helps search engines understand what people are looking for and gives better search results.

- It compares two different LLMs: **GPT-3.5 Turbo** and **Llama 2 7B**.
- It tests different ways to improve search results, such as **expanding queries** (adding more related words to a search to make it clearer).
- It works in multiple languages, including **English, German, French, Spanish, and Chinese**.
- It uses a technique called **Short Chain-of-Thought prompting**, which helps break down complex questions into simple steps.

I found this model impressive because it makes searching smarter and works across different languages, helping people worldwide find better information.

---

3. How the Model Works

This model is currently performing Information Retrieval (IR) in only five languages: **English, German, French, Spanish, and Chinese**. It improves search engines by making them smarter at understanding what a person is asking. Normally, search engines like Google just match words in a query with words in documents (LFIDF) and ChatGPT was working majorly on English based Information Retrieval but now through this model it also retrieves information in 5 different languages, but this model does much more:

1. **Expanding Queries**: If someone searches for "best laptops," the model adds more details like "top-rated laptops in 2024 with long battery life." With the help of Short-CoT.

2. **Using Smart Prompts**: It asks itself smart questions to understand a search better.
3. **Working in Different Languages**: It can rewrite searches in multiple languages, helping people get the best answers in their own language.
4. **Learning from Different Topics**: It improves itself over time by understanding different fields, such as medicine or finance.

---

4. Improvements Made and Their Effects

The Previous model was made even better by adding these improvements:

- **LLMs for search expansion** instead of just matching words.
- **Short Chain-of-Thought prompts** to make search queries more detailed.
- **Supporting multiple languages**, Doing Information Retrieval in Multiple language.

How These Changes Helped:

10% better recall (finding more useful results).

Previous vs. Current Recall Performance (LLM was not used Previously)

Language	Previous Recall	Current GPT-3.5 Turbo Recall	Current Llama 2 Recall
German	0.5724	0.7225	0.6099
English	0.8190	0.9027	0.8890
Spanish	0.7018	0.7770	0.7409
French	0.6528	0.7643	0.6089
Chinese	0.5599	0.6929	0.5007

Recall Improvement

Language	GPT-3.5 Turbo Improvement	Llama 2 Improvement
German	+26.3%	+6.5%
English	+10.2%	+8.5%
Spanish	+10.7%	+5.6%
French	+17.1%	-6.7% (Worse)
Chinese	+23.7%	-10.6% (Worse)

---

5. My Plan to Make It Even Better

Now that I understand this model, I want to work on improving it in the following ways:

1. **Adding More Languages**: Right now, the model works in **five languages (English, German, French, Spanish, and Chinese)**. I plan to add support for **more languages**, such as Hindi, Arabic, Japanese, Russian, and others, so that more people around the world can use it effectively.
2. **Making It Faster and Smaller**: Right now, this model takes a lot of computing power. I will study ways to **reduce its size and make it work faster**.
3. **Better Understanding of Small Languages**: Some languages don't have much data online. I will look into ways to **train this model to work better with low-resource languages**.

---

6. Conclusion: What I Learned

Through my research, I learned how different AI models work together to improve the way computers understand and process information. I now understand:

- **LLMs improve search engines** by making them smarter.
- **Expanding queries** helps get better results.
- **Multi-language support** can help more people find useful information.
- And I am Working on how to **further improve this model** to make it even more efficient and useful.

Reference:

GitHub - jantiegges/Query-Expansion-with-LLMs: Multilingual Information Retrieval with Large Language Model-Driven Query Expansion

---

Submitted by:
Riya (221210088)

Sahil Dhillon (221210092)