Knowledge graphs and LLMs

Importance of Knowledge Graphs

The algorithms which are used in the creation of Gen AI applications have always been present but used with different architectures like Transformers , GANs and VAEs but the difference that sets apart the old from new is : the enormous parameters utilized in the training of the model, the size and variation of the datasets how different algorithms are combined for more efficiency.

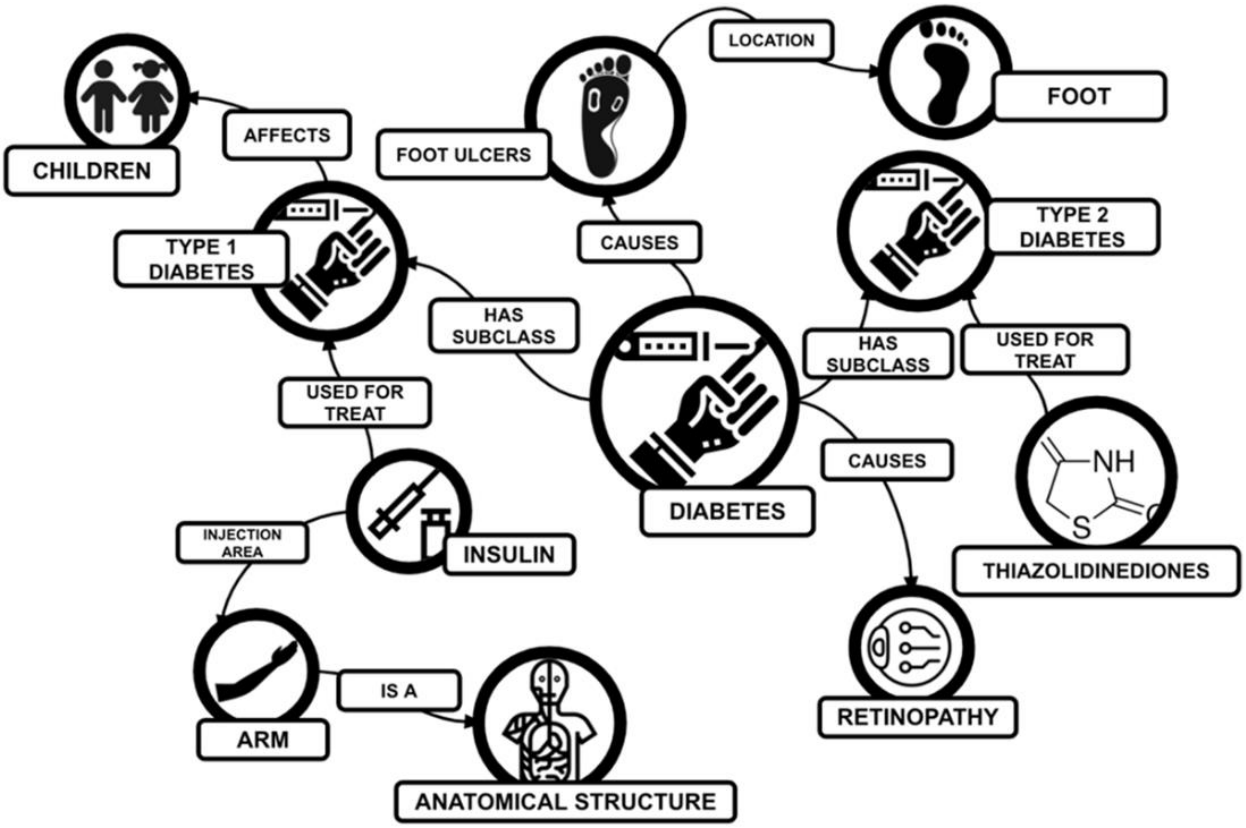
But even with newer changes and applied algorithms there are shortcomings when observed in the field on biomedicine, law, etc., where we need specific domain knowledge, high accuracy in the responses and contextuality in the explanation to be provided which are the practical limitations to be faced and these are due to the black box nature of LLM that dont access the factual knowledge, training is done on outdated generic datasets thus lacking information on specific areas to be concetrated and the lack of contextual and relatetional information which becomes crucial to acheiving more robust, trustworthy and explainable predictions.

These limitations can be overcome by knowledge graphs which equip the model with relational data, experiences, environmental characteristics, cultural aspects and social norms leading to the development in such as way where in those specific domains human life is relevant and where up-to-date information and contextual knowledge in specific domains are not only important but also crucial to acheive required goals.

Thus implementing Knowledge graphs together with llms can make up for the limitations to be faced by llm making them a good combination specifically in the area of natural language modelling and understanding that empower/develop the intelligent models used for advisory purposes for critical situations more reliable.

Knowledge Graphs

KG allows for the representation of human knowledge into machines in a structured graphical manner allowing them to behave intelligently where, Nodes in the graphs are real-world entities, the lines connecting these nodes are relationship between these nodes, Properties can be associated with these nodes and relationships.



A sample Knowledge graph in medical field

From the figure knowledge graph represents concepts and connections explicitly, they are stored in the form of nodes and relationships on top which it is possible to reach good responses which are reasonable and useful in inference. This type of knowledge system is most useful for the empowering complex intelligent systems capable of accomplishing complex tasks and reaching important decisions.

Even though KG have various advantages in helping for the development of the intelligent systems there have been certain challenges for their implementation like the cost, time, effort to build/maintain them, intricate access patterns required to navigate across multiple nodes over the graph, structure of the results with information scattered across multiple nodes and relationships.

Building KG is a crucial task and it requires recognizing and extracting heterogenous data relevant to the nodes and connections within the KG from external sources which is either structured or unstructured. Dealing with structured data is easier than unstructured, but even the process of integrating complex graphs into a common type graph which makes it even more controllable and fully predictable. Unstructured data is difficult to extract information from due to the reasons being typos, multiple languages, pronoun misspellings, Different styles in writing.

These factors contribute to the complexity of processing human languages, necessitating advanced techniques and robust systems to accurately extract and interpret information from unstructured text.

Once a KG has been built we can use its interrelated knowledge for the purpose gaining great insights and good support for our intelligent systems. While also giving flexibility in defining the schema for KG when dealing with heterogenous data and handling complex connections in the domain. But for those who do not know how to access KG it is difficult to give out a query and use the most out of the KG, even though it is possible to give upfront query and analysis to build specific intelligent systems but it limits the types of users and support given by the system.'

One of the best use-cases of KG is google search which enhances the user search experience by giving out the most relevance user queries as this approach gives out not only strings but emphasizes in the search of things showing the importance of KG in this application of searching. But this is also limiting the capabilities of KG and has not been used to its fullest extent which could be utilized for various purposes.

KG contain a plethora of information and analysts use them for answering specific domain queries but there is complexity rise due to the graph querying process and need for specific user interfaces limiting its potential to small user base, the future where we domain specific knowledge important individual use natural languages for querying and the intelligent system could find the correct answers by querying the graph effectively, potentially transforming the results into quick, simple summaries is where we want to reach using KGs.