**Ideas.**

-Deepti Saravanan

1. **Semantic Ambiguity – NLU**

* Regulations from different documents must not be contradictory. Otherwise, the regulations could become ambiguous semantically.
* NLU models could be used to capture contradicting regulations for semantic ambiguity analysis.

1. **QnA model – Source Expansion**

* Regulatory documents – seed corpora
* Based on the answer regulation, additional information retrieved from related documents connected to it in the SEBI knowledge graph.
* The retrieved additional information from multiple documents are summarized and provided with the answer regulation and answer span.
* The user can also view the entire document via the highlighted interactive KB.
* **Refinement**: The retrieval model (currently using classical tf-idf based model) could be updated with SOTA ideas like dual-embeddings to rank the retrieved relevant regulations.

1. **Link Prediction (cross-country relation)**

* Utilizing the SEBI knowledge graph as training data, models could be used to predict links between regulations from different countries (in the form of bipartite graph).
* These linked regulations could be further analyzed for differences.
* Using temporal information (amendments) of these linked regulations to study the market behavior of countries.
* Link prediction could also use temporal information instead of traditional edge weighing method.

1. **Amendment prediction – Social Media Analysis**

* Social media mining of posts related to amendments (Eg using hashtags from twitter).
* Sentiment Analysis – Higher the negativity -> more probable to get amended.

1. **SEBI vs Outsider interpretations**

* News Clarifications data could be used to analyze the interpretation differences.
* This sample of regulations could be analyzed for similarities and other exploration.