**Existing Systems**

**1)"Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks" (Lewis et al., 2020)**

Large pre-trained language models have been shown to store factual knowledge in their parameters, and achieve state-of-the-art results when fine-tuned on downstream NLP tasks. However, their ability to access and precisely manipulate knowledge is still limited, and hence on knowledge-intensive tasks, their performance lags behind task-specific architectures. Additionally, providing provenance for their decisions and updating their world knowledge remain open research problems. Pretrained models with a differentiable access mechanism to explicit non-parametric memory have so far been only investigated for extractive downstream tasks. We explore a general-purpose fine-tuning recipe for retrieval-augmented generation (RAG) — models which combine pre-trained parametric and non-parametric memory for language generation. We introduce RAG models where the parametric memory is a pre-trained seq2seq model and the non-parametric memory is a dense vector index of Wikipedia, accessed with a pre-trained neural retriever. We compare two RAG formulations, one which conditions on the same retrieved passages across the whole generated sequence, and another which can use different passages per token. We fine-tune and evaluate our models on a wide range of knowledgeintensive NLP tasks and set the state of the art on three open domain QA tasks, outperforming parametric seq2seq models and task-specific retrieve-and-extract architectures. For language generation tasks, we find that RAG models generate more specific, diverse and factual language than a state-of-the-art parametric-only seq2seq baseline.

**2)Can LLMs be Good Financial Advisors?: An Initial Study in Personal Decision Making for Optimized Outcomes**

Increasingly powerful Large Language Model (LLM) based chatbots, like ChatGPT and Bard, are becoming available to users that have the potential to revolutionize the quality of decision-making achieved by the public. In this context, we set out to investigate how such systems perform in the personal finance domain, where financial inclusion has been an overarching stated aim of banks for decades. We asked 13 questions representing banking products in personal finance: bank account, credit card and certificate of deposits and their inter-product interactions, and decisions related to high-value purchases, payment of bank dues, and investment advice, and in different dialects and languages (English, African American Vernacular English, and Telugu). We find that although the outputs of the chatbots are fluent and plausible, there are still critical gaps in providing accurate and reliable financial information using LLM-based chatbots.