# Yubo Wang (王钰博)


ywangnx@connect.ust.hk | Wyb0627 | Yubo Wang
ywangnx.student.ust.hk | wyb--0627 | +86 13756022978


## Research Interest

Large Language Models, Retrieval Augmented Generation, LLM Agent.

## Education

2021 – Present **Ph.D. Student/Candidate, CSE Department, HKUST**
Supervised by Prof. Lei Chen.
Work closely with Dr. Haoyang Li, Dr. Shimin Di and Dr. Hao Xin

2019 – 2021 **M.Sc. Computer Science, CISE Department, University of Florida**.
Supervised by Prof. Eric Jing Du

2015 – 2019 **B.Sc. Computer Science and Technology, CCST Department, Jilin University**.
Supervised by Prof. Fengfeng Zhou

## Selected Publications

- **AGRAG: Advanced Graph-based Retrieval-Augmented Generation for LLMs (ICDE 26 (CCF-A), 1st Author)**
  Design a semantic guided graph expansion algorithm based on Steiner Tree, to solve the information summarization and creative generation queries.
- **GORAG: Graph-based Online Retrieval Augmented Generation for Dynamic Few-shot Social Media Text Classification (WWW 26 (CCF-A), 1st Author)**
  Models document-label correlations as weighted graph edges with only a few labeled data; dynamically updates the graph with new nodes and edge weights to incorporate streaming trustworthy data.
- **KGLink: A Column Type Annotation Method that Combines Knowledge Graph and Pre-trained Language Model (ICDE 24 (CCF-A), 1st Author)**
  Leverages tabular structures to filter and refine retrieved Knowledge Graph (KG) information. Carry out end-to-end training for tabular understanding tasks.
- **Enhancing LLM Multi-hop Reasoning with Knowledge Graph-Guided Verifiable Dense RL Rewards (Ongoing, 1st Author)**
  KG for reasoning data synthesis and dense, verifiable reward function design for GRPO post-training.
- **Understanding the Embedding Methods on Hyper-relational Knowledge Graph (CIKM 25 (CCF-B), 1st Author)**
- **MGRAG: A Graph-based Multimodal Retrieval-augmented Generation System with Large Language Models (In Submission, 1st Author)**
- **ClinicalAgents: Multi-Agent Orchestration for Clinical Decision Making with Dual-Memory (In Submission, 3rd Author)**

## Industrial Projects

- **Adaptive Zero-shot Document Classification System.**
  *Webank, 2024.9-2025.9*
  Designed a graph-based RAG system, formulated the document-label correlation as weighted edges in a 0-shot manner, enabling the graph retrieval of query-related labels to aid LLM classification.