Fanyi Kong

♦Kong.fa@northeastern.edu ♦ https://kfy123.github.io/kfyswebsite/ ♦ www.linkedin.com/in/fanyi-k-27903b186

Education

Northeastern University | **M.Sc. Industrial Engineering** (Sep 2018 – May 2020 | Bos, MA)

- Relevant Courses: Data Mining(A), Deep Learning(A-), Computation and Visualization(A-)

China University of Labor Relations | **B.Eng. Safety Engineering** (Sep 2012 – Jun 2016 | Beijing, China)

- Relevant Courses: Epidemiological Statistical Analysis, Epidemiological Statistical Analysis, Basic Medicine

Honors and Awards: Scholarship Recipient (Top 5%), Merit Student (Top 3%), Led Robotics Training Group: 3rd Place (Fifth Tsinghua University Undergraduate Engineering Training Competition), Team Leader of a project funded by the Beijing University Students' Scientific Research and Entrepreneurship Program

Major Publications

- Zeng, H.†, Kong, F.†(2023). ***Active Learning: Encoder-Decoder-Outlayer and Vector Space Diversification Sampling***. *(†: Co-first authors)*. Mathematics, MDPI4353,Mathematics 2023, 11(13), 2819; https://doi.org/10.3390/math11132819
- Kong, F; Liu,H.;Li,S.;Ngai, Ka; Yu,Z. (2024). ***Predicting drug-drug interactions using heterogeneous graph neural networks: HGNN-DDI***. Accpted by Applied and Computational Engineering. Note: I, along with co-author Ka Yan Ngai, decided to withdraw from this publication after discovering that two team members outsourced the code for their contributions. https://drive.google.com/file/d/1rLpgExCsNWXwkyl2KrGEGGpVXQwvBLwI/view?usp=share_link
- Kong, F; Yu S. (2024). ***Measuring Common and Unique Information for Multimodal Medical Classification***. Submitted and under review at the Conference on ICCV. https://drive.google.com/file/d/1qVLYhWCR3S3GNwv_Sauxxad7zbFDzihH/view?usp=drive_link

Research Experience

Measuring Common and Unique Information for Multimodal Medical Classification

Multimodal Medical Classification, Information Theory (Aug 2024 - Nov 2024 | Online)

- Developed the Common and Unique Latent Representation (COULAR) framework to extract shared and modality-specific information from diverse medical data, enhancing classification performance and interpretability.
- Introduced an information-theoretic approach based on Gács-Körner common information for decomposing latent representations, improving feature disentanglement and reducing redundancy.
- Designed an autoencoder-based architecture to separate common and modality-specific information, preserving complementary insights while increasing diagnostic clarity.
- Achieved superior accuracy, F1 score, and AUC on three real-world datasets (BRCA, ROSMAP, LGG), outperforming state-of-the-art multimodal fusion techniques.
- Demonstrated the framework's ability to handle complex, high-dimensional medical data and improve clinical diagnostic capabilities.

Predicting Drug-Drug Interactions using Heterogeneous Graph Neural Networks: HGNN-DDI

GNN, Drug-drug Interaction, Machine Learning (Jul 2023 – Aug 2023 | Shanghai, China)

- Collaborated with Shanghai Jiao Tong University to develop a graph neural network (GNN) model with attention mechanisms for drug-drug interaction (DDI) predictions, using heterogeneous graphs where nodes represent drugs and proteins, and edges represent DDIs, drug-protein interactions (DPIs), and protein-protein interactions (PPIs).
- Curated and processed datasets, including 192,284 DDIs from DrugBank and over 4 million relationships from PrimeKG, integrating clinical descriptors and mapping drug-protein relationships to enhance model inputs. Preprocessed molecular data using ChemBERTa and calculated SMILES-based chemical structure similarities with RDkit.
- Designed a robust GNN architecture combined with a Multi-Layer Perceptron (MLP) for multi-class DDI predictions, achieving 96.86% accuracy by combining topological and chemical structure information.
- Optimized model performance with cross-validation and hyperparameter tuning (dropout, batch normalization), addressing challenges such as promiscuously interacting drugs and surpassing traditional machine learning methods like Decision Trees.
- Visualized feature spaces with t-SNE to analyze data clustering and interpret model behavior, generating healthcare-relevant predictions with associated probabilities to assess the risks and benefits of drug combinations.

Active Learning: Encoder-Decoder-Outlayer & Vector Space Diversification Sampling

Neural Network, Vector Space Diversification (Jan 2023 – Mar 2023 | Shanghai, China)

- Developed a machine learning pipeline under the supervision of Prof. David Woodruff (CMU), integrating an Encoder-Decoder-Outlier (EDO) framework and Vector Space Diversification (VSD) Sampling to enhance data diversity, minimizing labeling efforts, and optimizing GPU memory with buffer-stored encodings.

- Implemented the pipeline using a pre-trained "all-mpnet-base-v2" Sentence-BERT encoder and a 3-layer ResNet trained with the Nadam optimizer, achieving nearly linear time complexity in sampling and enabling efficient parallelization.
- Validated the approach through experiments, demonstrating superior performance on small datasets by reducing overfitting and achieving high accuracy while significantly lowering computational and labor costs.

Predicting In-Hospital Mortality of ICU Patients by Machine Learning

Feature Engineering, Model Diversity, Boosting Algorithm **(Feb 2019 - Apr 2019| Boston, MA)**

- Searched related materials and cleaned 299,264*42 datasets using Python. Used PFE to reduce the data dimension. Selected better features to train the model.
- Built models with two measures, unsupervised model Markov Chain and supervised models Random Forest, SVM, ANN.
- Applied three boosting algorithms (GBR, LightGBM, XGBoost) to improve output after bagging models.

Work Experience

Nanjing FiberHome Technology Co., Ltd.

Algorithm Engineer **(Dec 2020 – Apr 2023 | Nanjing, China)**

English Speech Recognition

- Led the development of the company's first English ASR product, achieving accuracy comparable to leading systems.
- Crawled over 60,000 hours of labeled open-source English audio and collected 600 Chinese-accented recordings to diversify the dataset and enhance recognition accuracy.
- Implemented SpecAugment for data augmentation, prepared datasets and labels using Python, and built a transformer-based model optimized through hyperparameter tuning and early stopping, achieving a Word Error Rate (WER) of 0.163.
- Tested, monitored, and packaged the model for deployment, ensuring robust performance and seamless integration.

Mandarin Speech Recognition

- Improved Mandarin ASR accuracy by 0.124 over previous versions through enhanced training methods and refined datasets.
- Prepared datasets and labels, developed a phoneme dictionary, and applied SpecAugment, Noise Addition, and Time Shift for robust model training.
- Built and optimized a Conformer-based model by fine-tuning hyperparameters and applying layer normalization to prevent overfitting, ensuring strong generalization and performance on new data.

Cantonese and Uyghur Speech Recognition

- Increased Cantonese accuracy by 0.163 and Uyghur accuracy by 0.137 over older versions using advanced training and refined datasets.
- IInvestigated linguistic features—such as Cantonese pronunciation rules and Uyghur sentence structures—to refine model inputs. Using Python, created a Cantonese pinyin dictionary and a Uyghur phonetic dictionary, and synthesized 70,000 Cantonese audio samples with varied timbres via the Ali interface to enhance the model's training quality.
- Prepared and processed datasets by screening missing characters, converting audio formats (from ".flac" to ".wav"), and generating new labels, ensuring comprehensive and high-quality training data.
- Built and trained DeepSpeech-based models for both languages, optimizing parameters based on gradient descent observations and achieving WERs of 0.167 (Cantonese) and 0.062 (Uyghur), representing significant accuracy improvements for both languages.

Multi-channel Speech Recognition

Second-place winner in the "2021 China Hualu Cup Data Lake Algorithm Competition"

- Developed a high-performance multi-channel speech recognition system using advanced preprocessing and modeling techniques.
- Compiled a high-quality dataset with diverse voice data to cover various speech patterns and accents, and developed an automated labeling system to streamline supervised training.
- Simulated real-world scenarios by merging and weighting voice data, applying beamforming to enhance directional signals, improving source separation and recognition accuracy.

Skills

- **Programming Languages**: Python, Shell Scripting, Batch Script, SQL, R, Java, SAS, Julia, Lingo
- **Operating System & Tools:** Linux, Windows, PyCharm, Git, Docker, Tableau, Jupyter
- **Framework:** PyTorch, TensorFlow, ESPnet, Mozilla DeepSpeech, Wav2Letter++, Kaldi, Numpy, Pandas
- **Cloud Service:** Amazon Web Services (AWS), Alibaba Cloud, Tencent Cloud
- **Technical Skills:** AI, Neural Network, Deep Learning, Machine Learning, Algorithm Design, Data Mining