Yufei Li

(+1) 469-371-4598 | yli927@ucr.edu | yul091.github.io/liyufei.me/ | yul091 | yufei-ethan-li

Research Interests

- Language processing and generation, uncertainty estimation, intelligent and interactive systems
- **Research goal**: effective, efficient, and explainable AI with knowledge and interactions

Education

University of California, Riverside — *Riverside, CA*

Ph.D. in Electrical Engineering, Advisor: Cong Liu, Intelligence System Group — *2022 - Present*

University of Texas at Dallas — *Dallas, TX*

Ph.D. in Computer Science, Advisor: Cong Liu & Wei Yang, Intelligence System Group — *2020 - 2022*

University of California, San Diego — *San Diego, CA*

M.S. in Electrical Engineering, Thesis: Application of A* and Genetic algorithm in TSP path planning — *2018 - 2020*

Xi'an Jiaotong University — *Xi'an, China*

B.S. in Mechanical Engineering, Thesis: Aircraft fuel tank oil sloshing simulation based on SPH method — *2014 - 2018*

Work Experience

NEC Laboratories America, Inc. — *Princeton, NJ*

Research Intern, PyTorch — *May 2022 - Aug 2022*

- Annotated fields (name entities) in raw system log datasets and applied prompt-based seq2seq BART for field extraction
- Defined hierarchical relationships between fields and configured dynamic attributed graphs for feature representations
- Detected abnormal relationships in log messages using a temporal-attentive Graph-Transformer encoder

NEC Laboratories America, Inc. — *Princeton, NJ*

Research Intern, PyTorch — *Jun 2021 - Aug 2021*

- Annotated name entities and relations with regular expression rules on CVE texts for distant supervision
- Incorporated the pre-trained GPT-2 backbone into a sequence labeling framework for joint entity & relation extraction
- Proposed a bootstrap training procedure for denoising distant labels and selecting high-quality instances

SeekTruth Scientific and Technical Corporation — *Beijing, China*

Research Intern, TensorFlow — *Jul 2019 - Sep 2019*

- Built a joint key point & pose recognition model for character detection, and refined an adaptive discrimination definition model
- Designed a lightweight CNN for real-time calibration of online video frame directions

Publications

Distantly-Supervised Joint Entity and Relation Extraction with Noise-Robust Learning

Yufei Li, Xiao Yu, Yanchi Liu, Haifeng Chen, Wei Yang, Cong Liu | *In Progress, 2022*

MIMONet: Multi-Input Multi-Output On-Device Deep Learning

Zexin Li, Xiaoxi He, Yufei Li, Wei Yang, Lothar Thiele, Cong Liu | *In Progress, 2022*

SHARE: a System for Hierarchical Assistive Recipe Editing

Shuyang Li, Yufei Li, Jianmo Ni, Julian McAuley | *EMNLP 2022* [pdf]

Estimating Predictive Uncertainty Under Program Data Distribution Shift

Yufei Li, Simin Chen, Cong Liu, Wei Yang | *Preprint 2021* [pdf]

GLIB: Towards Automated Test Oracle for Graphically-Rich Applications

Ke Chen*, Yufei Li*, Yingfeng Chen., Changjie Fan, Zhipeng Hu, Wei Yang | *ESEC/FSE 2021* [pdf]

Projects

Enhance the Reliability of Deep Networks on Out-of-distribution Data — *Dallas, TX*

NLP & CV, PyTorch — *Sep 2021 - May 2022*

- Perturbing the text inputs from unknown distribution via synonym substitutions to yield trustworthy predictions
- Boosting the CNN model confidence of test inputs through label-preserving image transformations

Assessing the Reusability of Pre-trained Code Embeddings

NLP & SE, PyTorch

Dallas, TX

Sep 2020 - May 2021

- Proposed a low-cost offline metric for evaluating the generalizability of code embedding in SE downstream tasks
- Patched the generalizability of existing pre-trained embedding based on the semantic metamorphic relationship

Rethink Negative Sampling in Bayesian Personalized Ranking

Data Mining & Recommender Systems, PyTorch

San Diego, CA

Nov 2019 - Jun 2020

- Analyzed one limitation of the popularity-based sampling scheme in terms of non-uniform negative sampling bias
- Corrected the bias and designed related negative sampling distributions to boost the Bayesian personalized ranking (BPR)

Automatic Delivery Vehicle Design

Path Planning, Python & MATLAB

San Diego, CA

Mar 2019 - Jun 2019

- Incorporated the Courier and TSP travel agent problems into designing autonomous delivery vehicles
- Designed a path planning algorithm by incorporating the A* heuristic rules into the genetic evolution

How Cute is My Pet? Pet Adoption Speed Prediction

NLP & CV, TensorFlow

San Diego, CA

Jan 2019 - Mar 2019

- Parsed and tokenized the description text of pets, encoded the semantic features with Bi-LSTM
- Designed a joint framework to fuse the image feature, semantic feature, and other properties for pet popularity prediction

Honors & Awards

VEX Robotics International Competitions

Path Planning, C/C++

Louisville, KY

2016-2017

- Excellent award and runner-up in VEX Robotics World Championship (RECF) 2017, Louisville, KY, US.
- Excellent award and runner-up in VEX Robotics Asia Open 2016, Beijing, China.
- First-class honor in VEX Robotics China Open 2016, Xi'an, China.

National encouragement scholarship

Personal, top10 of 300+ applicants.

Xi'an, China

2014-2017

Skills

Language	Python, C/C++, JAVA, Bash, HTML, Markdown, MATLAB
Machine Learning	PyTorch, TensorFlow, Keras, scikit-learn
Miscellaneous	LaTeX, Git, SQL, Ansys, SOLIDWORKS, AutoCAD, Photoshop