Zemin Yang

E-mail: csyzm@cug.edu.cn ✳ *Telephone number:* +86-18122780892

Education

Bachelor's degree in Data Science and Big Data Technology *September 2020 - Now*
China University of Geosciences (Wuhan)
GPA: 90.08 (3.88/5.00)
Language Proficiency: CET-6: 498; CET-4: 582;

Research and project experience

Autonomous Driving Lane Detection Project Based on CLRNet *July 2023 - Now*
The Summer Camp for Outstanding Undergraduates in the School of Software *Zhejiang University*

- Deployed the CLRNet-related Docker environment on ModelArts of Huawei Cloud and replicated it for verifying its performance on TuSimple and other datasets, and used the pre-trained model to tune its performance on an internal dataset (Lane line dataset at the port of NingBo).
- Conducted a thorough study of current 3D vision-based monocular camera lane line detection solutions (including 3D-LaneNet, Gen-LaneNet, 3D-LaneNet+, CLGo, PersFormer, SALAD, CurveFormer, WS-3D-Lane, Anchor3DLane, BEV-LaneNet) and summarized the possible development and improvement directions of lane line detection models.

Project Risk Assessment based on AutoMLPRA *June 2023 - July 2023*
The RedBird Challenge Camp *Hong Kong University of Science and Technology (Guangzhou)*

- Developed a comprehensive project risk evaluation standard system by combining previous risk evaluation index systems, analyzing results of international engineering cases, and utilizing web crawling technology for data collection.
- Constructed an AI risk assessment model called Automated Machine Learning for Project Risk Assessment (AutoMLPRA) based on AutoML principles, consisting of NLP tools for feature extraction, Algorithm Selection module, Preprocessing module, Model Training & Tuning module, Explanation module, and Model Evaluation module.

Deep learning-based classification of remote sensing image data *February 2023 - June 2023*
Course Project *China University of Geosciences (Wuhan)*

- Deployed MindSpore environment on Huawei Cloud ModelArts and built DenesNet (DenseNet-121/169/201) models and ResNet (ResNet-50/101/152) models from scratch upon it.
- The tuning of the DenseNet and ResNet models using the strategies of Kaiming initialization and both gradual warmup and cosine decay for learning rate ended up successfully improving the Top1 accuracy of the models on the RESISC45 dataset.

Data mining analysis project based on credit card overdue data *May 2023 - June 2023*
Course Project, Kaggle competition *China University of Geosciences (Wuhan)*

- Deployed Toad, HDFS, Spark, and AutoML environments distributed on Deepin OS.
- Used the Toad repository for data preprocessing (EDA analysis, decision tree binning, feature filtering based on IV/Pairplot/Forward Stepwise Regression).
- Established machine learning algorithms (including Decision Tree, Random Forest, LightGBM, XGBoost, CatBoost, etc.) to build a Risk Assessment Classification Model using PySpark and AutoML, and ended up with an AUC value of 0.867467 on the top of the list.

Honors and Awards

First Prize in the 13th Mathercup Mathematical Contest in Modeling	*Jun 2023*
Meritorious Winner in the American Collegiate Mathematical Contest in Modeling	*Feb 2023*
Third Prize in the MathorCup MCM (Big Data Competition)	*Jan 2023*
Ranking:99(Top4%) in the CCF Big Data and Computational Intelligence Competition	*Nov 2022*
First Prize(Hubei Region) in the China Undergraduate Mathematical Contest in Modeling	*Oct 2022*
Second Prize in the 13th Mathercup Mathematical Contest in Modeling	*Jun 2022*
Second Prize in the 10th Teddy Cup Data Mining Competition	*Jun 2022*

Technical skills

Programming Languages/Tools	C++, Matlab, Python, Pytorch, MindSpore, Sklearn, Hadoop, HBase, Spark, Flink, Linux OS(Ubuntu, Deepin)
Other Software/Tools	MS-Office software(Word, Excel, PowerPoint, Visio), LaTeX, Adobe software(PS, LR, PR, AE, AI), SPSSPRO, DxO Nik, Edrawsoft