Park, Sungnam

15, Gwangpyeong-ro 19-gil, Gangnam-gu, Seoul, Republic of Korea
+821048708573 |sungnam1108@naver.com | angrypark.github.io

OBJECTIVE

To obtain an internship opportunity in data science, utilizing my computer programming skills and engaging in new challenges.

TECHNICAL SKILLS

- Python
  - Data Handling : Pandas, Numpy
  - Machine Learning : Scikit-learn, Statsmodels, H2O, Xgboost, Catboost, StackNet(java)
  - Deep Learning : Keras, Pytorch
  - Computer Vision : Opencv
  - Natural Language Processing : Genism
  - Data Visualization : Matplotlib, Plot.ly, Seaborn, Cufflinks
- SAS
  - Data Handling : SAS Imstat
  - Data Analysis : SAS Enterprise Miner
  - Data Visualization : SAS Visual Analytics
- Development Tools : AWS, Github

ACCOMPLISHMENTS

Certification Prize in Big Data Contest **2017**
Sponsored by Mirae Asset Daewoo

- Forecasted KOSPI200 with RNN and Xgboost
- Made stock selling strategy and resulted 178% profit in 5 years

Club Representative of Yonsei Big Data Club (YBIGTA) **2016.12 ~ 2017.6**
Sponsored by Naver D2

- Arranged industry-academic cooperation project with social dating company, Amanda
- Hosted 2 big data conferences, one is united conference with other big data clubs in Seoul

1st Prize in Big Data Conference **2016**
Held by Yonsei Big Data Club (YBIGTA)

- Recommended movies using recommender system and collaborative filtering
- Vectorized users' information and rating histories

EDUCATIONS

B.S. Industrial Engineering, **Yonsei University** (Seoul, Korea) **2013.2 ~ 2019.7**(expected)
Jungsan High School (Seoul, Korea) **2010 ~ 2012**

EXPERIENCES

Military Service **2014.2 ~ 2015.11**
Scatterlab – Machine Learning Intern **2018.3 ~ present**

PROJECTS

Fibers Defects Detection – Hyosung ITX

- Classified if fibers are defected using opencv and Scikit-image

Loan Repayment Forecasting

- Binomial classification(90,216 columns) with stacking various models (Xgboost, Adaboost, Catboost, Logistic Regression, Neural Networks)
- Tried feature engineering, making derived variables, parameter tuning and model ensembling
- Handled imbalanced data using ROC Curve and f1-score and tried oversampling, undersampling

Contents Recommendation

- Used 417,414 contents history data of LG U+ to recommend contents to users
- Tried to make implicit ratings using k-NN, SVD

Deep Learning using Generative Adversarial Networks

- Read papers of GAN and posted summaries to my blog (GAN, DCGAN, Disco-GAN)
- Implemented neural style transfer and 1-d gaussian distribution generation.
- Posts : https://angrypark.github.io/blog/

Sentiment Analysis

- Crawled 2,600 reviews of landmarks in Seoul in TripAdvisor
- Used Word2Vec and LDA to find emotion (Positive / Negative) of landmark

Human Resource Data Analysis

- Used Kaggle dataset of human resource data to predict if an employee will leave company or not
- Used Adaboost, Xgboost, Keras and made business strategy with predictions

Recipe Recommendation(RecipeLab)

- Crawled korean food recipes and embedded to vectors using Word2Vec and Doc2Vec
- Recommended recipes for food a user can cook with ingredients in fridge

Face and Deep Learning(FaceBigta)

- Implemented VGGNet, neural style transfer, GAN and applied to facebook profile image
- Tried to make a face fancy with transferring style similar to celebrities' faces