Wei-Ying Wang, Ph.D.

weiwang@wayfair.com   wayinone.github.io   +1-401-556-9066

Melrose, MA (work remotely)

SUMMARY

- Seeking a data scientist / Machine Learning Engineer position in which I would contribute to the success of a business

- Data Schience Tech Lead in Wayfair, specialized in product matching and classification.

- 5 years of experience of scalable solutions and ML pipeline construction

- Applied Mathematics Ph.D.

TECHNICAL SKILLS

Machine Learning	Classification, Regression, Deep Learning, NLP (Fasttext, CRF)
Programming	Python(numpy, keras, sklearn, pandas, airflow, kedro, pyspark)
Database	MSSQL, Vertica, Hive, Big Query

PROFESSIONAL EXPERIENCES AT WAYFAIR LLC.

Product Match

· Matching 60M Wayfair's products against 1.1B crawled products in the market

· Discovering a business rule by analyzing previous matched data, which find 85% of new matches (8M of match pairs) in 30 min, where previous method will take 3 years

· Developing a 2-staged machine learning model for white-labeled products that has recall of 95% at 20% false positive rate with Spark ML and Scala

· Constructing a scalable ML matching pipeline (for white-label products) that can finish entire product matching process in 6 days

· Matching model is regularly retrained with automated evaluation process to ensure safe deployment

· Increasing existing matches by 300% and enabling analysis of market share, selection gap, and pricing

Product Classification

· Classifying entire catalog (60M Wayfair products and 1.1B crawled competitor products) into 800+ Wayfair classes within 2 hours

· Achieving 90% accuracy with deep learning model

· Automated retraining and evaluating to prevent model degradation

Manufacturer Normalization

· Normalizing different manufacturer synonyms, e.g. "HP", "Hewlett-Packard", and "Hewlett Packard", are alias of the same manufacturer and should be normalized together

· Normalizing 129K distinctive manufacturer names into 10K manufacturers, which covers 97% of crawled data

· Clean and accurate result impacting many aspects company-wide, fueling analysis like product gaps, MSRP estimation, and product matching projects

Part Number Extraction

· Extracting part numbers from 130M competitor's product name and description

· Utilizing a conditional random field model to achieve 95% precision

Optimal Threshold Determination with Bayesian Methods

- Estimating match pair suggestions accuracy by the feedback look from human validation, and utilize Bayesian statistics to obtain a robust result
- Automatically choosing optimal threshold for different classes of products, and improve overall accuracy by 3%, corresponding to 30K+ hours of labor saving

Code Standardization
- Modifying an open-source template (kedro) to standardize team's code to ensure production grade of coding from begining, allieviating Data Scientists the burden of common infrastrcuture burden like Jupyter Notebook, Spark environment, and code testing.
- The first team in Wayfair to advocate the benefit of the code standardization effort with end to end execution
- Speed up develop to production velocity by 200% when launching a new ML pipeline, compared to similar projects carried by the other team
- Enabling the separation of Engineering and data science work, leading to the maximum efficiency of the team

EDUCATION

Brown University, Providence, RI   Sep 2010 - May 2017
Ph.D. Applied Mathematics (GPA: 3.9/4.0)
- Dissertation: Image Compression and Data Clustering: New Takes on Some Old Problems
- Advisor: Stuart Geman

National Taiwan University, Taiwan   Sep 2004 - May 2006
M.Sc. Mathematics/Track of Statistics (GPA: 3.8/4.0)

National Taiwan University, Taiwan   Sep 2000 - May 2004
B.A. Economics (GPA: 3.8/4.0)

EMPLOYMENT

Wayfair, Massachusetts   Sep 2017 - Now
Data Science Tech Lead
- Leading a team of 2 data scientist and 1 engineer, creating production pipelines related to product matching

Brown University, Providence, RI   Jun 2017 - Present
Postdoc
- Finding high dimensional manifolds in time series data

Academia Sinica, Institute of Mathematics, Taiwan   Nov 2008 - Aug 2010
Research Assistant

Military Service, Taiwan   Jan 2007 - Jan 2008
Coastal Patrol Corporal