


Ren Zhang

<https://ryanzhang.info>  [k in](#)

Belmont, MA 617-893-5778

ryanzjlib@gmail.com

EXPERIENCE

Nielsen - BASES Innovation

Senior Data Scientist

Needham, MA

Apr. 2019 - Present

- † **Demand Advisor:** Graph neural network model trained to predict how the consumer would rank a set of products. The model can be used to generate recommendations for the product to be more favorable over the competition. Researched on the model architecture, implemented and trained model in Tensorflow2.0, exposed as REST APIs.
- † **Attribute Based Forecasting:** Gradient Boosting Tree models trained to forecast the sales volume for new product launches. The solution allows clients to experiment and optimize product attributes combinations in a product line. Performed data retrieval, implemented data processing pipeline, automated hyperparameters tuning.
- † **Trend Detection:** Solution to associate retail data to product feature groups and then project feature groups into the trend and growth space. The tool helps clients make sense and identify important product features. Proposed and experimented changes to the existing clustering methods, automated the pipeline.

IBM - Watson Health

Data Scientist

Cambridge, MA

Apr. 2017 - Mar. 2019

- † **Explorys:** Big data platform for storing longitudinal EHR data with analytic tools on top. Designed Avro data schema for new record types, implemented the ETL pipeline. Created a React web app for displaying and editing Avro data schemas.
- † **Watson Health Cloud:** HIPPA and GXP compliant cloud solution. Worked on developing an internal dashboard aggregates and displays the status of applications running on the cloud.
- † **Watson for Genomics:** Research tool uses NLP and knowledge base to link clinical trials to the genomic sequence. Implemented graph traversal strategies on knowledge graph to enhance entity extraction results. Trained word vectors and recurrent neural network models to classify documents into genetic mutation types.

EDUCATION

Bentley University

Master of Science in Business Analytics / Data Science - GPA: 4.00

Waltham, MA

Sept. 2015 - Dec. 2016

- † **Coursework:** Programming, Databases, Hadoop, Time Series, Optimization, Statistical Learning, Data Mining
- † **Research Assistant:** Worked with Prof. Babaian on developing D3 interactive data visualization for graph datasets.

PROJECTS

- DCASE 2019 Freesound Audio Tagging Challenge:** 19th place solution using convolutional neural networks to tag audio clips. Models were trained with SpecAugment and Mixup to enhance generalizability and combat label noise.
- Recursion Cellular Image Classification:** 30th place solution using convolutional neural networks to classify microscope cell images. Incorporated ArcFace loss from metric learning and performed iterative pseudo label refinement.
- Jigsaw Unintended Bias in Toxicity Classification:** Silver medal solution fine-tuning BERT to classify internet comments. Applied layer-wise learning rate decay and mixed-precision training to efficiently tuning big models.
- KDD Cup 2018:** 30th place solution forecasting air quality indices of the future 48 hours for locations in Beijing and London. Experimented with neural networks in time series forecasting.

PUBLICATION

Zhou L., Zhang R., Chakraborty P., Farooq F., Hensley Alford S., Comparative Effectiveness of Edoxaban and Warfarin in Prevention of Stroke and Systemic Embolism in Non-valvular Atrial Fibrillation Using Observational Healthcare Data, in ISPOR 2018.

Babaian T., Zhang R., Lucas W. DTMi – a New Interface for Informed Navigation, in proceedings HCI International 2017.

SKILLS

Experties: Deep Learning, Machine Learning, Statistical Analysis, Optimization, Software Development

Languages: Python, R, Javascript, Java, Scala, SQL

Libraries: Pytorch, Tensorflow, cuDF, PySpark, Sklearn, Pandas, Numpy, Scipy, Flask, React, Matplotlib, Seaborn, D3

Systems: Linux, Shell Script, Git, Docker, Jenkins, Hadoop, Spark