Chen, Tianle

Contact: tianlechen@gmail.com
Website: tianle91.github.io

Experience

- **Rubikloud Technologies Inc.** *Data Scientist* (May 2019 - Current)
  - Building demand forecasting models for retailers to balance supply chain considerations and promotion campaign effectiveness.
  - Developed a general representation for promotion mechanics in order to produce accurate forecasts and to compare among different promotion types.
  - Implemented cloud-based feature extraction, training and inference pipelines on `GCP` with `PySpark` on `Dataproc` and `Docker`-based virtual machines.
- **TMX Group** *Research Intern* (Jan 2019 - Apr 2019)
  - Developed a generative model for latent representation of market states and dynamics.
  - Clustering on latent representations reveals the market features that are useful in producing intuitive understanding of market state predictions.
  - Implemented an efficient algorithm in `PySpark` on `Zeppelin` to extract high-resolution features from orders and trade tables to capture detailed information regarding order book state and flows.
- **Rubikloud Technologies Inc.** *Data Science Research Intern* (May 2017 - Dec 2018)
  - Developed a novel individualized demand forecasting model in `Tensorflow` and `Keras` for joint predictions for purchase arrival times over multiple products using a Recurrent Neural Network (LSTM).
  - The model adapts log-likelihood losses which exploit partial information to obtain accurate and flexible predictions, beating state-of-the-art machine learning approaches with minimal hyper-parameter tuning.
- **University of Toronto** *Student Researcher - Insurance Reserving* (May-Aug 2016)
  - Implemented a Cox Process with random arrival intensities for Incurred but Not Yet Reported Claims (IBNR) in `R`.
  - Observed arrival intensities are represented as true arrival intensities thinned by a reporting delay distribution. True arrival intensities and reporting delay are then estimated from observations.
  - Showed that this model is much better able to predict the number of unreported claims compared to aggregate models.

Publications

- Badescu A.L., **Chen T.**, Lin S., Tang D., *A Marked Cox model for the Number of IBNR Claims: Estimation and Application*, 2019, ASTIN Bulletin, Volume 49, Issue 3, pp. 709-739. `https://doi-org.myaccess.library.utoronto.ca/10.1017/asb.2019.15`
- **Chen T.**, Keng B., Moreno J., *Multivariate Arrival Times with Recurrent Neural Networks for Personalized Demand Forecasting*, 2018, Published in Proceedings of IEEE ICDM 2018 DMS Workshop. `https://arxiv.org/abs/1812.11444`

Education

- **University of Toronto** *PhD Statistics, Withdrew from Program* (2017 - 2019)
  - Coursework includes topics in Statistical Learning Theory such as PAC learning, Online learning and Boosting.
  - Received NSERC Engage (value of 25,000 CAD) and Mitacs Accelerate (value of 15,000 CAD) funding from 2017 to 2018 for research in demand forecasting at Rubikloud Technologies Inc.
  - Received Mitacs Accelerate (value of 10,000 CAD) funding in 2019 for research in market forecasting at TMX Group Inc.
- **University of Toronto** *MSc Statistics, GPA: 3.80/4.00* (2016 - 2017)
  - Coursework includes topics in Applied Statistics, Measure Theory and Machine Learning.
  - Performed teaching and grading duties for Probability, Multivariate Data Analysis and Statistical Methods for Machine Learning.
  - Awarded Ontario Graduate Scholarship - Masters (value of 15,000 CAD).
- **University of Toronto** *Hons BSc Statistics, GPA: 3.83/4.00* (2012 - 2016)
  - Awarded Dean's List (top 20^{th} percentile) for Years 2, 3, 4.
  - Awarded the Morneau Shepell Scholarship in Actuarial Science in Year 2 (value of 2,500 CAD) for coursework in Actuarial Science.
  - Received Undergraduate Student Research Awards (NSERC) in Years 3, 4 (value of 6,000 CAD each) for research in Loss Models and Reserving.

Technical Skills

- Proficient in implementing and developing *neural networks* using `Tensorflow, Keras`.
- Experienced in *machine learning* packages such as `scikit-learn, XGBoost, LightGBM`.
- Proficient with *programming languages* such as `Python, R` as well as *database management and distributed computing frameworks* such as `Pandas`, `SQL` and `Spark`.
- Experienced with *version control tools* such as `Git` and operating in `Linux` environments.

Last Updated: January 12, 2020