

# JIRI STODULKA

[www.linkedin.com/in/jiristodulka/](https://www.linkedin.com/in/jiristodulka/)

• [www.jiristodulka.com](http://www.jiristodulka.com) • [jiri@jiristodulka.com](mailto:jiri@jiristodulka.com) • <https://github.com/jiristo> • (647) 290-5049 • Toronto, M4E 2W9

## SOFTWARE & SKILLS

**Python:** Machine learning (Scikit-Learn), Deep Learning (Keras, PyTorch) recommender system (Surprise), visualization (Matplotlib, Plotly, Folium). **R:** Machine learning (regression), statistics, text-sentiment analysis, visualization (ggplot), time-series (ARIMAX), latent class analysis (e1071), MySQL (DBI). **Cloud Service:** AWS (SageMaker, EC2, S3), IBM Watson (NLP: Personality Insights)

## WORK HISTORY

### Data Scientist, BECLOUDREADY Inc. 12/2019 – present (Toronto)

- Applied ensemble methods, e.g. random forest regressor (Scikit-Learn) and deep learning (Keras), and forecasted demand for 4,000 items in client's inventory while reducing old forecasting error of 25% to 6%.
- Engineered 300 features and interaction variables from three available time-series. Leveraged ML techniques to select the most important features accounting for 85%  $R^2$  and 6% forecasting error.

### Customer Facing Data Scientist, SALESCHOICE INC. 03/2019 – 11/2019 (Toronto)

- Utilized IBM Watson and Natural Language (NLP) toolkits for R&D. Used linguistic analytics to infer individual personality characteristics from digital communication.
- Coordinated with IBM on OCE grant R&D Project.
- Launched enterprise sales cycles with major clients (e.g. Stripe, AdRoll).
- Created a playbook to communicate the value of AI to these and other clients.

### Data Analyst, GEMINI EYE CLINIC 09/2016 – 07/2018 (Prague)

- Initiated the idea of implementing a franchise model for the clinic. Performed background research (including traveling to the Netherlands to talk to franchising expert) and built ML models to predict the effect of the strategy on profits.
- Helped put in place the new franchising model, which ultimately increased revenues by 15%.
- Worked with large datasets, e.g. wrote SQL queries in R Studio (DBI), utilized tidyverse.
- Utilized data visualization techniques and created dashboards; e.g. ggplot2 in R and Pivot Charts.

### Data Analyst (Project): BAUSCH AND LOMB 06/2016 – 08/2016 (Toronto)

- Translated data insights to outsourcing company; reduced loss on service operations by ~ 30%.
- Identified Canada-wide maintenance costs; improved utilization of the personnel by ~ 50 %.
- Cleaned, visualized, and analyzed data.

## DATA SCIENCE PORTFOLIO (Projects and GitHub: [www.jiristodulka.com](https://www.jiristodulka.com))

### Collaborative Filtering: Matrix Factorization Recommender System 2019

- Built matrix factorization model on utility matrix with 600 users, 10,000 items, and 100,000 ratings.
- Trained and evaluated SVD and NMF algorithms with Surprise package.
- Cross-validated the models and found ideal hyperparameters to minimize RMSE.
- Translated information about the users' history, recommendations, and metadata from various datasets.

### Toronto Crime and Folium 2019

- Received acknowledgement from the Toronto Police Department on LinkedIn.
- In Python, clustered and visualized criminal neighborhoods in interactive heatmap.
- Concluded there is 5% chance to be bodily harmed in assault.

### Supervised Machine Learning with Scikit-learn and Diabetes 2019

- Improved previous work by integrating cross-validation and hyperparameters in pipeline.
- In Python, utilized K-NN, logistic regression, decision trees, and random forest on diabetes dataset.
- Recommended to use recall as a less biased and more appropriate performance measurement.

### Web Scraping Reddit: Text (Sentiment) Analysis 2019

- Visualized sentiment in time and performed text analysis; e.g. identified causes of negative sentiment.
- In R, scraped Reddit forum: Cannabis Legalization in Canada.

### Other Projects

- Naïve Bayes and Fake News Detection (Accuracy 70% on 20,800 news dataset).
- Genetic Algorithm in Optimal Launch Sequence for Innovative Drugs (17% revenue improvement).

## EDUCATION

### MA in Applied Economics (Machine Learning, Data Science, Quantitative Research) 2018

Center for Economic Research and Graduate Education – Economics Institute (CERGE-EI)  
CERGE-EI is in top 5% of globally ranked Economics institutions; US-Chartered Degree - New York.  
**Elected as a Speaker at Graduation Ceremony**

### BSc in Business Administration (Statistics, Economics) 2017

State University of New York in Prague (UNYP)  
UNYP offers Bachelor's, Master's and PhD programs in partnership with US/European Universities.  
**Academic Recognition for outstanding GPA**