Artem I. Yankov

yankovai@umich.edu
URL: https://sites.google.com/a/umich.edu/yankovai/
GitHub: https://github.com/yankovai
(401) 206-9133

TECHNICAL SKILLS

Languages: Python, R, Fortran, bash, SQLite, Matlab, LaTeX, Apache Pig
Operating Systems: OS X, Unix, Windows
Applications and Libraries: Numpy, SciPy, Pandas, BeautifulSoup, ggplot2, Git, SciKit-Learn, Twitter Streaming API, Tableau, Dakota, Maple
Other Skills: Web scraping, linear/logistic regression, support vector machines, uncertainty quantification, numerical linear algebra, reduced order modeling

EXPERIENCE

Data Scientist Intern May 2014-present
Globys, Seattle, WA

- Developed a simulator for the usage behavior of mobile phone subscribers in a contextual marketing framework using machine learning and statistical data analysis.
- Researched and developed algorithms for predicting customer churn, including Bayesian Bandits and Hidden Markov Models.

Graduate Student Research Assistant July 2010-present
University of Michigan, Department of Nuclear Engineering, Ann Arbor, MI

- Applied statistical and machine learning techniques to analyze the phenomenon of fission gas release in nuclear fuel performance.
  - Involved folding together experimental data and computer simulation results to improve predictive capabilities of computer codes.
- Research in how uncertainties in reactor simulation code input parameters propagate to output predictions.
  - Extensive collaboration with researchers from Oak Ridge National Laboratory to develop algorithms for quantifying uncertainties in nuclear reactor simulators when faced with large numbers of correlated, stochastic inputs.
  - Results of research published in leading journal and awarded first prize at a major technical conference.
- Implemented numerical linear algebra routines into primary software used by Nuclear Regulatory Commission to simulate nuclear reactor accident scenarios.

PERSONAL PROJECTS

StackOverflow Query Tag Extraction

- Used Python to analyze over 6,000,000 StackOverflow queries and corresponding tags using MapReduce-type framework.
- Reformulated raw textual data to a useful form and placed into SQL database for further processing.
- Developed an algorithm to automatically predict new tags based on similarity to queries in training data set with over 80% accuracy.

NCAA College Basketball Prediction

- Scraped a decade's worth of college basketball data from sports-reference.com/cbb using BeautifulSoup and stored results in SQLite database.

- Analyzed data and developed a logistic regression model to predict the outcome of unplayed games with some 14% lift.
- Ranked all college basketball teams based on simulations of predictive model.
- Created a visualization of predictive model performance using Tableau.

Twitter User Cravings

- Used Twitter Streaming API to investigate cravings of Twitter users.
- Utilized Apache Pig to filter and process relevant data.
- Created a visualization application using Tableau to present results.

EDUCATION

University of Michigan — Ann Arbor, MI
Ph.D Nuclear Engineering and Radiological Sciences — Expected 2014

Rose-Hulman Institute of Technology — Terre Haute, IN
B.S. Mathematics — May, 2010
B.S. Physics
Minor: Computational Science
Clarence P. Sousley Award for demonstration of exceptional performance in the mathematical sciences.