

# B. Selin Tosun

LinkedIn: [in/b-selin-tosun](https://www.linkedin.com/in/b-selin-tosun)

Github: [/BanuSelinTosun](https://github.com/BanuSelinTosun)

Patents, Publications, Honors, and Awards: [banuselintosun.com](https://banuselintosun.com)

Seattle, WA

(612) 991-1606

[selin.tosun@gmail.com](mailto:selin.tosun@gmail.com)

*Data Scientist with 9+ years of experience in analytical problem solving*

## Technical Skills

**Expertise:** Python (pandas, NumPy, Seaborn, Scikit-learn, TensorFlow, Caffe2), SQL, MatLab, Mathematica, Machine Learning, Feature Engineering, Experimental Design, Statistical Analysis, Git, AWS, MS Azure

**Proficient:** Spark, Hadoop, QGIS, AI, Flask, HTML, JavaScript, CSS, IBM Watson Analytics, Docker, Horovod, Octave, Databricks

## Professional Experience

**Senior Data Scientist**, Neal Analytics & Microsoft Azure Storage, Bellevue, WA April 2018 – present

- Built & developed & optimized object detection models (achieved over 48 % AP) from various Tensorflow Model Zoo algorithms via tensor-board, RetinaNet, and Facebook Detectron in caffe2 environment utilizing Docker images
- Utilized existing data sets, and curated new data sets using ffmpeg and lblImage (images and annotations)
- Built a pipeline for end-user preferred trainings and model comparison in Average Precision and Inference time
- Developed a pipeline to shorten the training time by number of nodes using distributed training by Horovod

**Data Science Fellow**, Insight, Seattle, WA January 2018 – March 2018

- Built & developed [takeapic.online](https://takeapic.online), a Facial Expression Analyzer using Convolutional Neural Network (CNN) in Python (Keras-TensorFlow, training with GPU-Cuda) to improve social media experience
  - Built multi-classification CNN with Facial Expression Data Base of 55K+ images and achieved >98% accuracy

**Data Science Student**, Galvanize, Inc., Seattle, WA June 2017 – September 2017

- Built & developed [street-smart-realty.com](https://street-smart-realty.com), a Real-Estate housing price estimator using Python (pandas, NumPy, scikit-learn, matplotlib) and QGIS Grid-Search of Random Forest, Gradient Boosting, XG-Boost, SVM, Elastic Net
- Gathered data from various resources: King County, Seattle Public Schools, Great Schools, Zillow
  - Achieved model performance of 11.3 % median absolute percent error

**Senior Process Engineer**, Intel Corporation, Hillsboro, OR April 2015 – April 2017

- Developed new plasma etching processes for continuously evolving state-of-the-art transistors; optimized the process to increase yield by 10% in < 4 months
- Published a white paper on upgrading the reliability of etch tools to improve part lifetime

**Post-Doctoral Research Associate**, University of Washington, Seattle, WA July 2013 – February 2015

- Built a novel spectrometer to analyze and improve solar cell device efficiencies as a function of thickness
- Modeled solar cell device performance for accurate detections in MatLab

**Research Assistant**, University of Minnesota, Minneapolis, MN January 2009 – June 2013

- Performed experimental design producing 1000+ samples state-of-the-art thin film solar cell devices and compared with the base-line (1000+ samples) to identify significant differences in each trial in MatLab
- Improved lifetime of these solar cell devices from < 20 years to 40+ years: [2 US Patents](#), [7 peer-review articles](#)

## Education

**Certificate in Data Science** September 2017  
Galvanize Inc., Seattle, WA

**Certificate in SQL & Python Fundamentals** April 2017  
SOLO Learn Inc.

**Ph.D. in Chemical Engineering** June 2013  
University of Minnesota, Minneapolis, USA (In top 3 Graduate ChemE programs in US: MIT, UMN, CalTech)

- Doctoral Dissertation Fellowship: awarded to top 1% of the graduating Ph.D. students

**B.Sc. in Chemical Engineering**, *summa cum laude* June 2007

**B.Sc. in Material Science and Metallurgical Engineering**, *magna cum laude* June 2008  
Istanbul Technical University, Turkey