AHRIM HAN, PH.D.

✉ ahrimhan@gmail.com 📞 (949)910-8751 📍 Irvine, CA (Permanent US Resident / Green Card Holder) in ahrimhan ⦿ ahrimhan

SUMMARY

I am transitioning from academia to industry to work as a data scientist and machine learning scientist. Over the past 14 years, I have researched software engineering in computer science to improve software design and code quality using statistics and machine learning techniques. I have a special interest in finding business impacts by leveraging big data intelligence.

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST)   Feb. 2007 - Aug. 2013
Ph.D. Computer Science 2013

Korea Advanced Institute of Science and Technology (KAIST)   Sept. 2004 - Feb. 2007
M.S. Computer Science 2007

Sogang University   Feb. 2000 - Feb. 2004
B.E. Computer Science 2004

PROJECTS

Sentimental Analysis of Movie Reviews using a Deep Learning Neural Network (Springboard)   Mar. 2019 - July 2019
We have developed a deep learning model for identifying and classifying the sentiment expressed in texts of movie reviews as positive or negative using Python with Keras. We especially used a pre-trained embedding, Glove, for the Embedding layer in a Convolutional Neural Network model and discussed the overfitting.

Prediction of Scores for Public Schools in California (Springboard)   Oct. 2018 - June 2019
I applied many of the data science-related techniques for data wrangling, exploratory data analysis, data visualization, and machine learning modeling. I provided the prediction models using the regression and classification algorithms for finding the inferior schools that need help.

Efficient Refactoring Candidate Identification (Korea University)   Mar. 2008 - Apr. 2017
As a sole principal investigator, I developed **an efficient refactoring recommendation system** that helps software developers change code more easily. In order to increase the computational efficiency of the heavy computation when evaluating a large number of refactoring candidates, I developed the following methods.
- I developed a *fast refactoring candidate assessment method* using the graph-based approach and matrix computation. This sacrifices some degree of precision but **significantly reduces the computation complexity,** which can be helpful in analyzing large-scale software.
- I suggested the *two-phased search-based refactoring identification method* for the efficient search space exploration by predicting refactoring candidates with higher chances of quality improvement. Compared to the no-reduction approach, our approach could be **13.5 (max) times faster in time.**

Improvement of Change-Proneness Prediction (KAIST)   Mar. 2006 - Feb. 2010
I developed **a more accurate change-proneness prediction model** to help software developers to focus their efforts on more critical maintenance activities**.** I devised the **new behavioral dependency metrics** to capture the dynamic aspects of the program, and the model accuracy (R-square) using these metrics is an 8% increase over the model using only program structural metrics.

EXPERIENCE

Springboard, *Data Science Career Track - Student (Deep Learning Specialization)*   Oct. 2018 - Aug. 2019
- Mastering skills in Python, SQL, data wrangling, data visualization, exploratory data analysis, and machine learning and deep learning methods
- Performed two data science-related capstone projects: "**Prediction of Scores for Public Schools in California**" and "**Sentiment Analysis of Movie Reviews using a Deep Learning Neural Network**"

Korea University, *Research Professor*, Seoul, South Korea   Sept. 2013 - Apr. 2018
- Advised graduate students to develop research topics and conduct the experiments
- Awarded $158,000 in grants from the National Research Foundation of Korea for my research projects, **"Efficient Refactoring Candidate Identification"**

Korea Advanced Institute of Science and Technology (KAIST), *Graduate Researcher*, Daejeon, South Korea   Mar. 2005 - Aug. 2013
- Actively performed my research project "**Improvement of Change-Proneness Prediction**" and implemented program structure analysis and metric measurement tools using Java and Python
- Mentoring and Taught in Computer Science courses: "Introduction to JAVA programming" and "Capstone Projects in CS"

PUBLICATIONS

Two-phase Assessment Approach to Improve the Efficiency of Refactoring Identification, *Ah-Rim Han, Sungdeok Cha*
IEEE Transactions on Software Engineering (TSE), Vol 44, No. 10, pp. 1001 - 1023, Oct. 2018

An efficient approach to identify multiple and independent Move Method refactoring candidates, *Ah-Rim Han, Doo-Hwan Bae, Sungdeok Cha*
Information and Software Technology (IST), Vol. 59, pp. 53–66, Mar. 2015

SKILLS

DATA ANALYSIS, VISUALIZATION, AND MODELING: Pandas, Numpy, Scipy, Matplotlib, Pyplot, Seaborn, Scikit-Learn, Keras, Tensorflow, PySpark, NLTK
STATISTICAL METHODS AND MACHINE LEARNING: Classification, Regression, Clustering, Hypothesis Testing, Deep Learning
PROGRAMMING LANGUAGES AND TOOLS: Python, SQL, Java, R, Fortran, Assembly, Markdown, Latex, SPSS
RESEARCH AREAS: Software Engineering, Refactoring, Software Quality Improvement, Software Design Measurement, Natural Language Processing