

AHRIM HAN, PH.D.

✉ ahrimhan@gmail.com 🌐 ahrimhan.github.io ☎ (949)910-8751 📍 Irvine, CA (Green Card) in ahrimhan 🔄 ahrimhan

SUMMARY

I am transitioning from academia to industry to work as a data 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. Throughout my research, I focused on software-related problems. Now I have a special interest in solving user-related problems such as user preference prediction by leveraging big data intelligence.

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST) Ph.D. Computer Science 2013	Feb. 2007 to Aug. 2013
Korea Advanced Institute of Science and Technology (KAIST) M.S. Computer Science 2007	Sept. 2004 to Feb. 2007
Sogang University B.E. Computer Science 2004 Magna Cum Laude	Feb. 2000 to Feb. 2004

PROJECTS

Sentimental Analysis of Movie Reviews using a Deep Learning Neural Network (Springboard) We have developed a deep learning model to automatically classify movie reviews as positive or negative in Python with Keras. We especially identify and deal with overfitting and use a pre-trained embedding in a neural network model to improve accuracy.	Mar. 2019 to June 2019
Prediction of Scores for Public Schools in California (Springboard) 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.	Oct. 2018 to June 2019
Efficient Refactoring Candidate Identification (Korea University) 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. <ul style="list-style-type: none">I developed the <i>fast graph-based coupling metric</i> for calculating the effects of suggested refactorings based on 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 <i>two-phased search-based refactoring identification method</i> for the efficient search space exploration by predicting refactoring candidates with higher chances of quality improvement. Compared to the no-reduction approach, our approach is 2.6 (min) to 13.5 (max) times faster in time.	Mar. 2008 to Apr. 2017
Improvement of Change-Proneness Prediction (KAIST) I developed a more accurate change-proneness prediction model to help software developers to focus their efforts on more critical maintenance activities. I developed the new behavioral dependency metrics that capture the dynamic aspects of the program. Using dependency metrics in conjunction with existing program structural metrics, we built the regression model. The accuracy (R-square) of our model is increased 8% than the model using only structural metrics.	Mar. 2006 to Feb. 2010

EXPERIENCE

Springboard Data Science Career Track - Student <ul style="list-style-type: none">Mastering skills in Python, SQL, data wrangling, data visualization, exploratory data analysis, and machine learning methodsPerformed two data science related capstone projects: "Prediction of Scores for Public Schools in California" and "Sentimental Analysis of Movie Reviews using a Deep Learning Neural Network"	Oct. 2018 to Current
Korea University Research Professor <ul style="list-style-type: none">Advised graduate students to develop research topics and conduct the experimentsAwarded \$158,000 in grants from the National Research Foundation of Korea for my research projects, <i>"Efficient Refactoring Candidate Identification"</i>Published the research results in the top conferences and journals such as <i>IEEE Transactions on Software Engineering</i> (Google Scholar)	Sept. 2013 to Apr. 2018 Seoul, South Korea
Korea Advanced Institute of Science and Technology (KAIST) Graduate Researcher <ul style="list-style-type: none">Actively performed my research project "Improvement of Change-Proneness Prediction" (papers cited 64 times)Mentoring and Taught in Computer Science courses: "Introduction to JAVA programming" and "Capstone Projects in CS"Implemented program structure analysis and metric measurement tools using Java and Python	Mar. 2005 to Aug. 2013 Daejeon, South Korea
Peace Corps Intern Responsible for organizing the intranet web pages in the Technical Infrastructure and Support Team under the Chief Information Officer	Aug. 2004 to Oct. 2004 Washington, DC, USA

SKILLS

DATA ANALYSIS, VISUALIZATION, AND MODELING: Pandas, Numpy, Scipy, Matplotlib, Pyplot, Seaborn, Scikit-Learn, Keras, Tensorflow, PySpark
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