SUNG PIL MOON

17 Scarlett Ct., New City, NY
317-361-5563

Email: monspo1@gmail.com
Portfolio: monspo1.github.io/
LinkedIn: linkedin.com/in/sungpilmoon
Blog: bit.ly/nycdsablog-sungpilmoon

SUMMARY

- Aspiring Data Scientist with experience in software development for academic, practical, and research projects
- Seeking for a data scientist position, where I can utilize my skills to effectively contribute to your team
- Proficient in programming, machine learning, problem solving, data analysis and data visualization

EDUCATION

Indiana University, School of Informatics and Computing, Indianapolis, IN **May 2015**

- *Ph.D.*: Human-Computer Interaction (Overall GPA: 3.95 / 4.00)

Carnegie Mellon University, School of Computer Science, Pittsburgh, PA **August 2006**

- *M.S.*: Information Technology in eBusiness, (Overall GPA: 3.57 / 4.00)
- Scholarship Recipient: Ministry of Commerce, Industry and Energy of Korea

Soongsil University, Seoul, KOREA **February 2004**

- B.S.: Computer Science (Overall GPA: 3.74 / 4.30)

EXPERIENCE

Data Scientist Fellow, NYC Data Science Academy, New York, NY **Jan. - Mar. 2016**

- **DataScienceJobAnalyzer project**: implemented the interactive cluster and sentiment analyses of web-scrapped job data from indeed.com and dice.com using Python, R and Shiny Dashboard
- **DataScientistSalaryComparator project**: implemented interactive data scientist salary comparator against other eight professions in the US using R and Shiny Dashboard (based on 167,278 prevailing wage data from US Department of Labor)
- **Kaggle BNP Paribas Cardif competition project**: participated in the competition to predict category of user claims based on features available from a large data (30~40% missing and anonymized) in the early process using supervised learning methods.
- **DataSci4Good project**: implemented the improved visualization and embedded recommendations component to the Wise shiny app designed for young people to develop financial literacy utilizing Shiny, Python and several machine learning methods

Researcher and Research Engineer, School of Informatics, Indiana University, Indianapolis, IN **Sept. 2009 - May 2015**

- **Persuasive mobile infoVis app development** Dec.2012 – Dec.2014
  Developed a mobile application for sustainable motivation in a context of running (PhD dissertation topic) using elements of information visualization, gamification and social grouping for obese people to change sedentary lifestyle to more active.
- **Shared decision space infoVis tool for planning and analysis of large healthcare data** Sep.2012 – Dec.2013
  Developed a decision support tool providing shared decision space information visualization of large and complex patient data sets for intelligence analysts, health care professionals and patients. Funded by MITRE research corporation (www.mitre.org)
- **TopHealthTrends infoVis tool development** Nov.2011 – June 2012
  Developed an information Visualization tool showing local health-related Twitter trends to aid jobs of health-related experts
- **Bridging the Situation Space to Decision Space Gap** Oct.2009 – Sep.2011
  Developed a prototypical decision-making simulator to aid first emergency responder providing visualization of multiple decision options. Funded by MITRE research corporation (www.mitre.org)

Research and Teaching Assistant, Carnegie Mellon, ISRI, Pittsburgh, PA **Nov.2006 – May 2007**

- Led the "Robots to the Rescue" class and offered a guidance of introduction to robotics and programming to undergrad.
- Developed an interactive simulator with Microsoft Robotics Studio to communicate between four-wheeled robots and simulators

SKILLS

- **Programming**: Python (scikit-learn, numpy, scipy, pandas, matplotlib, seaborn, etc.), R (data manipulation & visualization: Shiny, ggplot, dplyr, knitr, caret, rpart, etc.), SQL, Hadoop (MapReduce, Hive), GitHub, Java, Javascript, HTML5 / CSS, Android Mobile, jQuery Mobile, Adobe Flex web / mobile programming (Actionscript)
- **Statistics and Machine Learning**: Linear and Logistic regression, Generalized linear models, Feature selection & engineering, K-Nearest Neighbors, Principal Component Analysis (PCA), Lasso & Ridge Regression, K-means clustering, Hierarchical clustering, Classification and Regression trees, Decision trees, Random forests, Support vector machines, Associations Rule, Naïve Bayes, Gradient Boosting Machines (GBM), XGBoost, Neural Networks, Time series models

PATENT

System and Method for Producing Video Map

- Application No: US20100077307 A1 (published on March 25, 2010)