|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Jamal Ching-Chuan Chen  陳慶全  Data Scientist   * **Contact**   (work) +886-963-855-707  (private) +886-966-676-326  Mail: [zw12356@gmail.com](mailto:zw12356@gmail.com)  GitHub: ChingChuan-Chen   * **Education**   **National Cheng Kung University, Tainan, Taiwan**  Master of Statistics  GPA: 4.0 / 4.0  **Thesis:**  A Classification Approach Based on Density Ratio Estimation with Subspace Projection   * **Languages**   Chinese – Naïve Speaker  English – Fluent  Japanese Intermediate (N3)   * **Skills**   R / MatLab – Master  Statistics – Advanced  Python – High-Intermediate  Bash / SQL – High-Intermediate  Machine Learning – Intermediate  C++ / Scala – Intermediate  C# / Java / JavaScript – Basic | * **Summary**   I am Jamal. I work as data scientist in a in multinational enterprise of wafer manufacturing. I am also…   * + A system developer with domain knowledge and strong technical skills.   + A skilled engineer in distributed computing, data preprocessing and data visualization.   + A programmer skilled with R, Python, Shell, MatLab, Scala, SQL and C++.   + A statistician worked deeply with theoretical or applied statistical methods.   + A machine learner acquainted with different algorithms to solve real problems. * **Work Experiences**  |  |  | | --- | --- | | **2018.09 – Present** | **Senior Data Scientist** | | **CIM Department, TSMC, Taichung** | | **2016.07 – 2018.08** | **Data Scientist** | | **CIM Department, TSMC, Taichung** | | **2015.09 – 2016.06** | **Research Assistant** | | **Institute of Statistical Science, Academia Sinica, Taipei** |  * **As a (Senior) Data Scientist in TSMC**  1. **The target in TSMC:**  * Develop automation systems on quality control of wafer processing from a big volume of data (3 billions per day).  1. **Highlights in TSMC:**  * Develop an algorithm to identify the defects on the wafer via a neural network model. * Develop an algorithm of process changing detection which suits for final WAT data. * Introduce a yield analyzing procedure to find out important variables. * Introduce a correlation visualizing system to reveal the relationships between measurements. * Build up a big data solution for our department. * Construct a developer-friendly environment for developing R and Python behind firewall. * **As a Research Assistant in Academia Sinica**  1. **The target in Academia Sinica**  * Complete at least one research in the field of functional data analysis.  1. **Highlights in Academia Sinica**  * Construct a scheduled task to digest data crawled from Taiwan freeway bureau. * Use functional clustering and functional regression to impute the missing values and predict future flow, occupancy rate and speed. * Build up an interactive visualization system to show the trends of flow, occupancy rate and speed. * **Journals**  1. Ping-Yang Chen, Ching-Chuan Chen, Chun-Hao Yang, Sheng-Mao Chang and Kuo-Jung Lee, milr: Multiple-Instance Logistic Regression with Lasso Penalty, The R Journal (2017), 9:1, P. 446-457  * **Awards**  1. Third Place, TSMC Kaggle Competition for the Defect Recognition (2017) 2. Honorable Mention, Competition for Data Analysis with R in Taiwan (2014) |