Salary Prediction Models

* Salary Prediction in Data Science Field Using Specialized Skills and Job Benefits – A Literature Review (2022) ([Link](https://jati.sites.apiit.edu.my/files/2022/07/Volume6_Issue3_Paper10_2022.pdf))
* Tee Zhen Quan, Mafas Raheem
* Salary prediction models specifically for the data science field.
* Data science roles are complex and require a lot of various skills so models should be based on that to improve accuracy
* Statistical Methods:
  + Linear and polynomial regression are good for understanding relationships between variables when there is linearity, predict salary based on factors like job level and experience but are limited in handling multiple complex variables accurately
* Machine Learning methods: Random Forest and K-Nearest Neighbors (KNN)
  + Random Forest: improved accuracy by combining multiple decision trees, which makes predictions more stable
  + KNN: Was effective with minimal tuning but showed some bias at different salary levels
* Deep Learning: Neural Networks and Bidirectional GRU with CNN for complex and unlabeled data
  + Valued skills first then predicted salary, this allowed feedback that improved accuracy
  + Neural network models performed better than traditional techniques but required more computation
* No gender-related issues or biases in salary prediction mentioned
* Salary Prediction Model using Principal Component Analysis and Deep Neural Network Algorithm (2023) ([Link](https://www.researchgate.net/publication/378207843_Salary_Prediction_Model_using_Principal_Component_Analysis_and_Deep_Neural_Network_Algorithm))
* Salary prediction model utilizing Principal Component Analysis (PCA)
  + Using Pca to reduce data dimensions this makes it easier for models to learn key patterns
  + PCA is used first to identify the most influential variables from the data.
  + Then the selected features are then used in a DNN, it is trained to classify salary levels
  + The DNN model outperformed other methods like Decision Trees and Random Forests, it achieved high accuracy with fewer prediction errors
* No gender-related issues or biases in salary prediction mentioned
* Statistical Machine Learning Regression Models for Salary Prediction Featuring Economy Wide Activities and Occupations (2022) ([Link](https://www.mdpi.com/2078-2489/13/10/495))
* Yasser T. Matbouli, M. Alghamdi
* Framework to predict salaries across different jobs and industries in Saudi Arabia based on economic activities and occupational groups
* Bayesian Gaussian Process Regression (GPR), Artificial Neural Networks (ANN):
  + Better Performance, models improved prediction accuracy and error rates compared to traditional multiple linear regression
* Mentions gender as part of the demographic data that can affect salary levels but avoids using it as a specific characteristic
* Income Prediction Using Machine Learning Techniques (2024) ([Link](https://escholarship.org/content/qt6d01c9v7/qt6d01c9v7_noSplash_d6307f10bef85009fad51d5837b90bb1.pdf?t=seap4w))
* Kahyun Jo
* Income prediction, mainly if someone earns more than $50,000
* Methods: Logistic Regression, Decision Trees, Random Forest, Neural Networks
  + Random Forests achieved the best results in terms of accuracy, precision, and recall
  + advanced tuning techniques were used to optimize models like Random Forests and Neural Networks
* Gender was taken into account, and paper mentions demographic factors might contribute to biases