

Time-Series Forecasting of Academic Careers

Dylan O'Driscoll

Information on the career of an academic researcher is readily available online. It can be difficult if not impossible for a human being evaluating all these statistics to predict how a researcher will perform in the future. This can make it difficult for universities to decide which researchers to hire for certain positions, or for researchers themselves to know where they stand in relation to their colleagues. This dissertation investigates different ways of predicting how the different aspects of an academic researcher's career can evolve over time, using machine learning and time series techniques. Using data gathered from a public repository, different algorithms and strategies were compared in order to predict the citation count, publication count and future universities of academic researchers. For predicting metrics, 3 different time periods were analysed using two different prediction strategies. The three time periods were predicting their total citation count 10 years after first publishing based on data from their first 5 years, predicting citation count at 15 years based on the first 10 and at 20 years based on the first 15. In one strategy a single prediction was used to attain these values, and in the other the values of citation count, publication count and h-index for each individual year were predicted one year at a time, with predicted values of previous years used to predict the values for the next year. Future universities a researcher may go on to work at are predicted based on the order of the universities they have worked at previously