DevFlair: A Framework to Automate the Pre-screening Process of Software Engineering Job Candidates

R.T.R Jayasekara[1], K.A.N.D Kudarachchi[1], K.G.S.S.K Kariyawasam[1], S.L Jayasinghe[1], Dilini Rajapaksha[2], and Samantha Thelijjagoda[3]

[1]Department of Computer Science and Software Engineering, Sri Lanka Institute of Information Technology, Sri Lanka.
[2]School of Computing Technologies, Royal Melbourne Institute of Technology, Australia.
[3]SLIIT Business School, Sri Lanka Institute of Information Technology, Sri Lanka.
it19129204@my.sliit.lk, it19121352@my.sliit.lk, it19147024@my.sliit.lk, it19126234@my.sliit.lk , dilini.rajapaksha@rmit.edu.au, samantha.t@sliit.lk

Abstract

The HR department of a technology company receives hundreds of job applications for each Software Engineering related vacancy. Evaluating a candidate by looking at the curriculum vitae may appear to be easy during the pre- screening process. However, an automated pre-screening process using Natural Language Processing and Machine Learning methodologies would help the recruiter to obtain a more accurate and deeper understanding of the candidate. In this paper we propose DevFlair; a framework to automate pre-screening Software Engineering job candidates. DevFlair uses data from social media, GitHub, and open-ended questionnaires to predict the Big-Five personality traits, analyze technical skill expertise, and analyze the experience in using industry-related online platforms. After analysis, the candidates are ranked according to their personality and technical skill levels. We conduct the personality prediction experiments using a social media posts dataset annotated with gold-standard Big- Five personality labels. We train FastText classification models and compare their accuracy against other state of the art classification models. The comparisons conclude that the FastText classification models substantially outperform the state of the art classification models when predicting Openness, Conscientiousness, and Agreeableness personality traits.

Keywords

Pre-Screening, Human Resources, Personality Prediction, Natural Language Processing, Machine Learning