

Counterfactual Explanations for Employment Services

Raphael Mazzine Barbosa de Oliveira, Sofie Goethals, Dieter Brughmans, David Martens



Abstract of the paper. This paper introduces the applications of a specific XAI methodology, Counterfactual Explanations, to the employability field. We show the usage through eight potential use cases..

Methodology

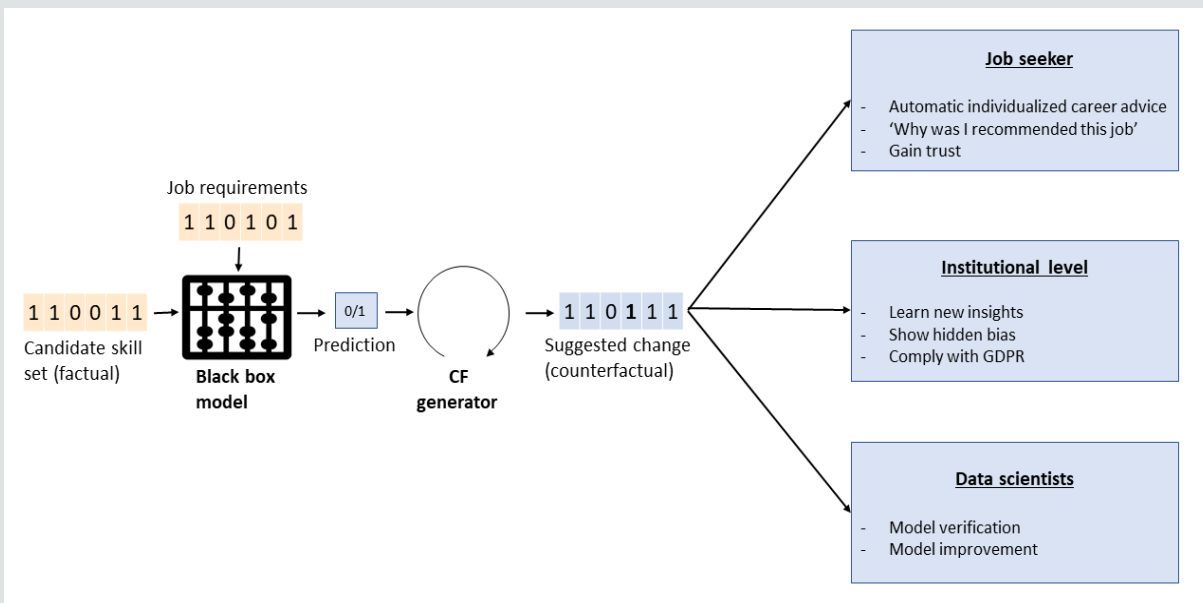
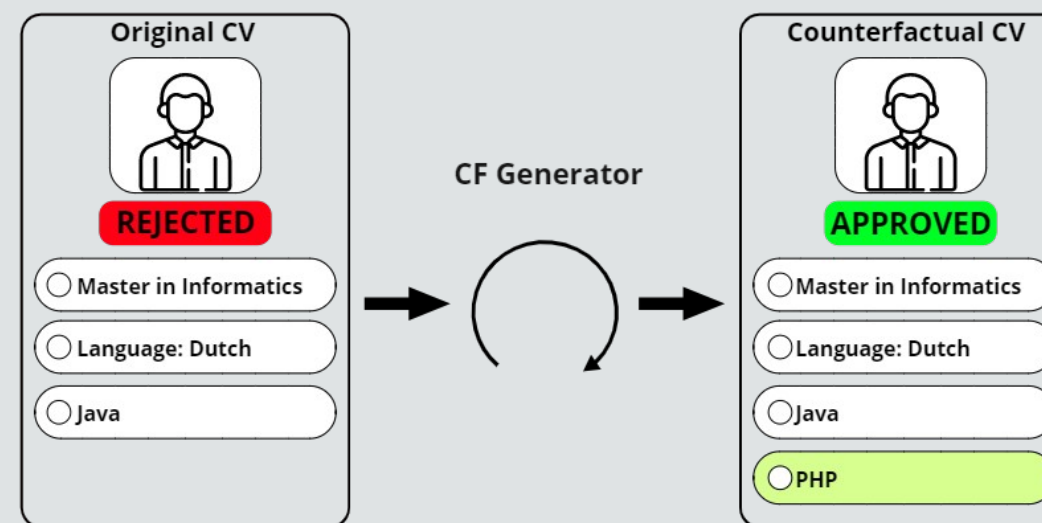


Fig. 1 – Schematic overview of used methodology

- ❖ Train any black box model to predict whether a person is applicable for a job.
- ❖ Use an algorithm to generate the counterfactual explanations (We used NICE/SEDC)

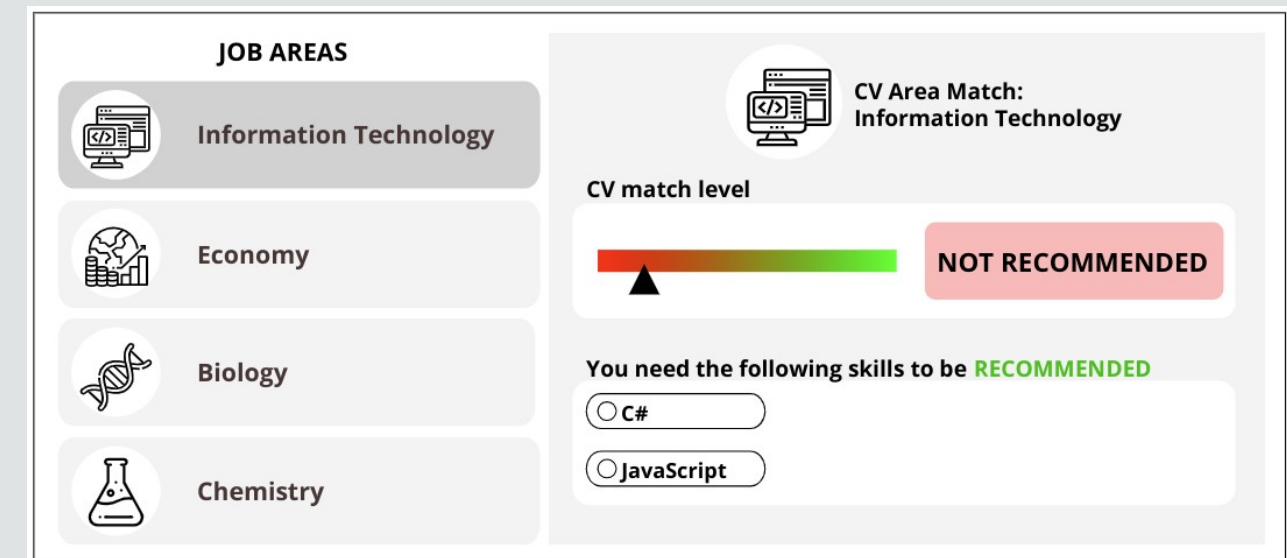
Counterfactual generation



Use cases

- ❖ Automatic individualized career advice
- ❖ Acquiring new insights from the model and data
- ❖ 'Why was I recommended this job?'
- ❖ Gain trust and social acceptance
- ❖ Show hidden bias
- ❖ Legal compliance
- ❖ Model improvement
- ❖ Model verification

Use case 1: Career advice



Use case 2: New insights

