

Analysis of Marie Skłodowska-Curie Actions (MSCA) evaluations and models for predicting the success of proposals

Ilaria Rodella¹, Andrea Sciandra², Arjuna Tuzzi³

¹ University of Padova – ilaria.rodella@unipd.it

² University of Padova – andrea.sciandra@unipd.it

³ University of Padova – arjuna.tuzzi@unipd.it

Abstract

This study analyses Evaluation Summary Reports (ESRs) of Marie Skłodowska-Curie Actions (MSCA) Individual and Postdoctoral Fellowships proposals at the University of Padua (Unipd), spanning Horizon 2020 and Horizon Europe from 2015 to 2022. The aim is to identify recurring strengths and weaknesses in the evaluation process, recognizing the most important and recurrent features of successful proposals. Nearly 400 ESRs were analysed by employing keyword extraction and correspondence analysis (CA) to map relationships between words and variables such as project success. While CA did not clearly distinguish between successful and unsuccessful proposals, machine learning was applied. The coordinates from CA were used to predict project outcomes. Comparisons were made with models using only textual features and those employing transformers, specifically, BERT contextualised embeddings. Results showed that using a Large Language Model (LLM) for text representation improved prediction accuracy compared to other methods. However, it highlighted challenges in interpretability and emphasised the need for explicable methods in the absence of words. Overall, the study provides valuable insights for refining support services and training at Unipd, highlighting the effectiveness of LLMs in prediction while acknowledging the interpretive challenges associated with their use.

Keywords: Marie Skłodowska-Curie Actions, evaluation, keyword extraction, correspondence analysis, large language model, text classification.