

# DATA SCIENCE ASSIGNMENT

## INTRODUCTION

This take-home assignment is designed to test your skills in data science and presentation. You will be presenting the results of your analysis in a virtual interview. Your final submission should include **A)** a pdf document containing your **code** in markdown and **B)** a pdf document containing **slides** highlighting your approach and results.

## SITUATION

A Polish client is providing insurance services in the context of bankrupt companies. The client would like you to predict the likelihood of companies to go bankrupt based on financial characteristics. To help you with your task, you should use historical data on company survival provided.

### 1. BANKRUPTCY PREDICTION USING FULL DATA

Suppose your task is to predict the chance of bankruptcy after 1, 2 and 3 years from the features in the data set. How would you approach building a model and which techniques do you consider? Please train a model and document how well it performs.

### 2. SELECTION OF IMPORTANT FEATURES

In reality, only a small number of features will be directly accessible, and gathering additional features is costly. How do you approach selecting the features with the highest predictive power for company bankruptcy? Which relations do you find?

### 3. NEXT STEPS

Let's say the client likes your analysis but would like you to come up with additional insights. Which next steps would you recommend in the analysis?

### 4. PRESENTATION

Please prepare a submission in the format discussed in the introduction. You should mention the techniques you considered and the thoughts behind your evaluation in a way that could be presented to the client. Ideally, your slides will serve as a starting point for further discussion of your analysis.