| | | Total |
|---|---|---|
| **Theme** | **Data preprocessing and visualization** | **20** |
| | Dataset is loaded and visualizations are created to understand the output and relationships between attributes and output | 5 |
| | Relevant variables are chosen and correlated features are eliminated | 5 |
| | Outliers are removed, if necessary | 5 |
| | New features are created, if necessary | 5 |
| **Theme** | **Dataset engineering** | **10** |
| | Dataset is split into training and test sets | 5 |
| | Standardization is applied, if necessary | 5 |
| **Theme** | **Model definition** | **30** |
| | A machine learning algorithm is chosen and fit to the data | 10 |
| | A neural network is fit to the data | 10 |
| | Hyperparameters are optimized using grid search, if necessary | 10 |
| **Theme** | **Model evaluation** | **30** |
| | Each model is evaluated on the test set | 10 |
| | Best model is chosen | 5 |
| | Visualizations of some predictions are made and compared to true predictions with the best model | 15 |
| **Theme** | **Code and documentation** | **20** |
| | Every block of code is commented and justified | 10 |
| | Methodology and choices are well-justified | 10 |
| **Theme** | **Overall performance** | **20** |
| | The chosen model is able to accurately predict the output variable | 10 |
| | The methodology used is thorough and well-documented | 10 |