You are provided with an anonymized dataset containing both categorical and numeric variables available when the claims were received by renowned insurance company. All string type variables are categorical. There are no ordinal variables.

The "target" column in the train set is the variable to predict. It is equal to 1 for claims suitable for an accelerated approval.

claims for which approval could be accelerated leading to faster payments 🡺 1

claims for which additional information is required before approval 🡺 0