THE MACHINE LEARNING CANVAS

Designed for: Designed by: Date:

Iteration: .

PREDICTION TASK

Type of task? Entity on which predictions are made? Possible outcomes? Wait time before observation?

Identificação e classificação da qualidade de vinhos tintos e brancos.

DECISIONS

How are predictions turned into proposed value for the end-user? Mention parameters of the process / application that does that.

Decidir se um vinho é considerado "bom" ou "ruim" com base em suas características físico-químicas.

VALUE PROPOSITION

Who is the end-user? What are their objectives? How will they benefit from the ML system? Mention workflow/interfaces.

Fornecer aos produtores de vinho uma ferramenta automatizada para avaliar a qualidade de seus produtos com base em características objetivas.

DATA COLLECTION

Strategy for initial train set & continuous update. Mention collection rate, holdout on production entities, cost/constraints to observe outcomes.

Coleta de dados sobre características físico-químicas de vinhos tintos e brancos.

DATA SOURCES

Where can we get (raw) information on entities and observed outcomes? Mention database tables, API methods, websites to scrape, etc.

Origem dos dados: "Modeling wine preferences by data mining from physicochemical properties" de P. Cortez et al. (2009).

IMPACT SIMULATION

Can models be deployed? Which test data to assess performance? Cost/gain values for (in)correct decisions? Fairness constraint?

Simular o impacto do modelo de classificação na avaliação da qualidade dos vinhos e fornecer informações sobre o desempenho do modelo.

MAKING PREDICTIONS

When do we make real-time / batch pred.? Time available for this + featurization + post-processing? Compute target?

Utilização de técnicas de Machine Learning, como Random Forest, XGBoost e Regressão Logística, para fazer previsões sobre a qualidade dos vinhos.

BUILDING MODELS

How many prod models are needed? When would we update? Time available for this (including featurization and analysis)?

Treinamento de modelos de classificação usando os algoritmos de Random Forest, XGBoost e Regressão Logística.

FEATURES

Input representations available at prediction time, extracted from raw data sources.

Variáveis preditoras:

- Acidez fixa: medida da concentração de ácidos fixos no vinho.
- Acidez volátil: medida da concentração de ácidos voláteis no vinho.
- Ácido cítrico: medida da concentração de ácido cítrico no vinho.
- Açúcar residual: quantidade de açúcar residual no vinho.
- Cloretos: quantidade de cloretos presentes no vinho.
- Dióxido de enxofre livre: quantidade de dióxido de enxofre livre no vinho.
- Dióxido de enxofre total: quantidade total de dióxido de enxofre no vinho.
- Densidade: medida da densidade do vinho.
- pH: medida do nível de acidez ou alcalinidade do vinho.
- Sulfatos: quantidade de sulfatos presentes no vinho.
- Álcool: teor alcoólico do vinho.

Variável alvo:

- Qualidade: avaliação da qualidade do vinho, representada por uma escala numérica de 0 a 10.

MONITORING

Metrics to quantify value creation and measure the ML system's impact in production (on end-users and business)?

Monitoramento contínuo do desempenho do modelo de machine learning e atualização dos dados de treinamento para manter a precisão das previsões.

OWNML.CO

ONLINE COURSE

Master the Machine Learning Canvas

Learn a step-by-step process to get to a complete and detailed Machine Learning Canvas. This will help you...

- Validate the feasibility of your ML use case ideas.
- Boost collaboration within your team.
- Anticipate issues that would otherwise come up during implementation or in production.

More details at **ownml.co/plan**