

Projeto de Ciências de Dados

Objetivo

Desenvolver pelo um modelo de ciências de dados (Algoritmo Genético, Fuzzy, Rede Neural, Deep Learning, Árvore de Decisão,...) para resolução do problema de previsão de utilização do cartão de crédito.

Contexto

Imagine que você foi contratado por uma empresa de cartão de crédito que a devido a pandemia teve a sua venda de novos cartões diminuídas, pois ficou difícil vender os cartões de crédito na rua, modelo tradicional, por isso o marketing da empresa teve a ideia de comprar uma base de novos clientes de uma empresa de empréstimos.

Sendo assim, você deve avaliar a base atual dos clientes da empresa e construir modelos que identifiquem o perfil do cliente que utiliza cartão de crédito e dos clientes que tem cartão de crédito qual é o possível valor gasto por cada novo cliente.

Dados levantados

- ID: Número único do cliente na base
- Renda: Valor da renda anual do cliente (Valor dividido por R\$ 10.000)
- Limite: Limite de crédito estabelecido para o cliente
- Rating: Índice de qualidade do empréstimo para os clientes (quanto maior o rating melhor pagador é o cliente)
- Cartões: Quantidade de cartões que o cliente possui
- Idade: Idade do cliente
- Educação: Números de anos de estudo do cliente
- Sexo: Masculino ou Feminino
- Estudante: Se o cliente estuda ou não
- Casado: Se o cliente é casado ou não
- Etnia: Qual é a etnia do cliente
- **Gasto Mensal: Valor Gasto com o cartão mensalmente**

Dicas

O arquivo **CCARDBALANCE.xlsx** possui vários atributos lembre que modelos do tipo rede neural ou deep learning utilizam apenas dados numéricos, dessa forma será necessário transformar algumas variáveis em números (ex.: Casado (Sim ou Não) em Casado (1 ou 0)).

Primeiro construa um modelo que faça a predição da coluna **Gasto Mensal** ele é o objetivo principal da compra de nova base de clientes.

Avaliação

Deverá ser entregue um relatório constando os passos para realização dos modelos construídos, coloque figuras, imagens e gráficos com os resultados. Não esqueça das métricas de avaliação como acurácia e erro médio.

Inclua os scripts e/ou códigos para avaliação.