

Mariana Costa

Cientista de Dados Sênior

São Paulo, SP

Com 6 anos de experiência em análise de dados e machine learning, especializada em projetos de impacto social e otimização de processos para empresas de varejo. Defensora da diversidade em tecnologia e cofundadora do grupo Mulheres em Dados Brasil.

HABILIDADES

Python



Scikit-learn



Power BI



Estatística



Pandas



SQL



TensorFlow



ETL



PROJETOS EM DESTAQUE

Previsão Vendas Express

Modelo de machine learning para previsão de demanda de produtos p supermercados Express, reduzindo o desperdício de alimentos. Lidera da coleta de dados à implantação.

Resultados do projeto

- Redução de 38% no desperdício de produtos perecíveis
- Aumento de 12% na margem de lucro da rede
- Economia anual de R\$ 2,7 milhões

Educação em Números

Plataforma pública de dados aberto que cruza informações de escola desempenho no ENEM para identificar desigualdades educacionais p dados do projeto.

Resultados do projeto

- Mais de 50 mil acessos nos primeiros 6 meses
- Utilização dos dados por 12 prefeituras para definir políticas de inves
- Prêmio de melhor projeto de dados abertos de 2023 pela Open Data

ARTIGOS

Como usar dados para reduzir desigualdades: lições do projeto Educação em Números

Quando falamos de ciência de dados, é comum associar a área a projetos ferramenta tem um potencial enorme para gerar impacto social. O projeto percepção, durante um trabalho voluntário em uma escola pública de São educação eram feitas sem base em dados concretos.

e resultados do ENEM por escola, tudo isso de forma anonimizada para proteger a identidade dos alunos. Para limpar e estruturar os dados, criamos visualizações interativas que mostram o desempenho do aluno e desempenho no ENEM, por bairro.

O resultado foi que 12 prefeituras passaram a utilizar nossa plataforma para monitorar o desempenho dos alunos, e conseguimos identificar que bairros com menor renda tinham até 10% de alunos em bairros de classe alta. Uma lição que aprendi nesse projeto é que dados são necessários para reduzir desigualdades, mas só geram impacto se forem apresentados de uma forma que todos os envolvidos possam entender e utilizar.

Machine learning para previsão de vendas no varejo: um guia prático

A previsão de vendas é uma das aplicações mais comuns de machine learning no varejo. Se você não sabe por onde começar. No projeto Previsão Vendas Express, desenvolvemos um modelo para prever as vendas de 120 lojas de uma rede de supermercados, e conseguimos reduzir o desperdício em 15%.

O primeiro passo para desenvolver um modelo de previsão de vendas é coletar dados históricos de vendas e fatores externos como feriados, eventos locais e até previsão do tempo, que impactam as vendas. Depois, é importante testar diferentes algoritmos, como regressão linear, árvores de decisão e redes neurais, para ver qual tem o melhor desempenho para o seu caso específico.

Uma dica importante é não tentar criar um modelo único para todas as lojas, mas sim criar modelos específicos para cada região, pois o comportamento de compra em bairros de classe alta é diferente dos bairros de classe baixa. O resultado foi que a precisão do modelo aumentou de 68% para 89%, o que permitiu reduzir o desperdício em 15% para a rede de supermercados.