

Aplicação de Machine Learning na Previsão de Entrega Logística

Bronstrup, Felipe César; Lourenzzi, Wagner Simões.

Machine Learning, Data Science, Pesquisa Operacional

O cenário das organizações é apresentado cada vez mais competitivo em todos os aspectos, sejam eles econômicos, onde cita-se os custos envolvidos aos processos das organizações, e de qualidade vinculados tanto aos produtos, quanto aos serviços prestados aos clientes. Segundo Da Silva (2018), os clientes estão cada vez mais exigentes e seletivos sobre os produtos e serviços oferecidos pelas empresas, e com a globalização a competitividade elevou o nível de exigência sobre as organizações, fazendo com que necessitem desenvolver estratégias de gestão e planejamento cada vez mais eficientes para o crescimento organizacional. Tendo em vista que o resultado contábil da organização objeto deste estudo ocorre por meio da entrega das mercadorias aos clientes, sendo portanto a variável *LeadTime* de entrega um importante fator decisório sobre os embarques prioritários, o objetivo deste trabalho está na aplicação de um modelo de *Machine Learning* através do algoritmo *Random Forest* para classificação dos clientes que terão entrega efetiva no mês vigente. O modelo foi elaborado na linguagem de programação *Python 3* e dos pacotes *Pandas* e *ScikitLearn* disponíveis na linguagem. O modelo é alimentado com os dados históricos das entregas realizadas pela organização no período de sete anos. Os registros de entrega totalizam 260 mil registros diários de entregas, sendo as variáveis preditoras utilizadas no modelo os atributos de Estado de Origem, Estado de Destino, Data de Faturamento, Volume de Mercadorias e Data de Faturamento. Os dados do modelo foram separados entre dados de treino e teste na proporção de 70% para treino e 30% para teste, foram também asseguradas iguais proporções de registros entregues e não entregue em ambos os conjuntos. O modelo treinado foi submetido aos dados de teste, resultado em um percentual de acerto de 89%, foi também gerada a matriz confusão das previsões do modelo, resultado em 39.793 registros positivos, 3.978 registros falso positivos, 28.452 registros negativos e 3.612 registros falso positivos. O modelo apresenta resultados satisfatórios para aplicação prática, sendo a sequência do estudo a utilização do modelo treinado para priorização dos clientes que devem ser faturados para melhoria dos resultados operacionais e financeiros da organização.