



Data Science Academy

www.datascienceacademy.com.br

Formação Cientista de Dados

Projeto com Feedback 8

Modelagem Preditiva em IoT - Previsão de
Uso de Energia

Este projeto de IoT tem como objetivo a criação de modelos preditivos para a previsão de consumo de energia de eletrodomésticos. Os dados utilizados incluem medições de sensores de temperatura e umidade de uma rede sem fio, previsão do tempo de uma estação de um aeroporto e uso de energia utilizada por luminárias.

Nesse projeto de aprendizado de máquina você deve realizar a filtragem de dados para remover parâmetros não-preditivos e selecionar os melhores recursos (melhores features) para previsão. O conjunto de dados foi coletado por um período de 10 minutos por cerca de 5 meses. As condições de temperatura e umidade da casa foram monitoradas com uma rede de sensores sem fio ZigBee.

Cada nó sem fio transmitia as condições de temperatura e umidade em torno de 3 min. Em seguida, a média dos dados foi calculada para períodos de 10 minutos. Os dados de energia foram registrados a cada 10 minutos com medidores de energia de barramento m. O tempo da estação meteorológica mais próxima do aeroporto (Aeroporto de Chievres, Bélgica) foi baixado de um conjunto de dados públicos do Reliable Prognosis (rp5.ru) e mesclado com os conjuntos de dados experimentais usando a coluna de data e hora. Duas variáveis aleatórias foram incluídas no conjunto de dados para testar os modelos de regressão e filtrar os atributos não preditivos (parâmetros).

Seu trabalho agora é construir um modelo preditivo que possa prever o consumo de energia com base nos dados de sensores IoT coletados. Recomendamos usar RandomForest para a seleção de atributos e SVM, Regressão Logística Multilinear ou Gradient Boosting para o modelo preditivo. Recomendamos ainda o uso da linguagem R. Os datasets de treino e de teste estão em anexo a este pdf.

Quando concluir o projeto, envie os scripts e datasets para projeto@dsacademy.com.br. Caso os datasets usados sejam muito grandes, armazene em um diretório virtual (existem vários na internet, como Google Drive ou Dropbox) e envie o link para que nossa equipe possa baixar os datasets. Se os arquivos foram pequenos (uma amostra do dataset original), envie no anexo junto com o script. Documente seu script tanto quanto possível.

Caso prefira, disponibilize seu projeto no Github e envie o link do seu repositório para nossa equipe no e-mail projeto@dsacademy.com.br. Nesse caso, o Readme do repositório deve constar que este trata-se de um projeto da Formação Cientista de Dados da Data Science Academy.