**Tratamento e previsão de variáveis climáticas INMET – versão 01**

**Breve Descrição**

O principal objetivo do programa é treinar e avaliar diferentes modelos de aprendizado de máquina, incluindo modelos tradicionais e modelos de séries temporais como o ARIMA, especificamente para dados de variáveis climáticas do INMET (Instituto Nacional de Meteorologia). Ele fornece funcionalidades para a criação de pipelines de processamento de dados, validação cruzada e cálculo de métricas de desempenho, como MAE, MSE, RMSE e MAPE.

**Funcionalidades específicas relevantes:**

* Criação de pipelines de pré-processamento de dados, incluindo tratamento de características numéricas e categóricas.
* Validação cruzada utilizando K-Fold para modelos tradicionais e TimeSeriesSplit para modelos de séries temporais.
* Cálculo e retorno de métricas de desempenho, incluindo erro absoluto médio (MAE), erro quadrático médio (MSE), raiz do erro quadrático médio (RMSE) e erro percentual absoluto médio (MAPE).
* Suporte a modelos de aprendizado de máquina como regressão linear, floresta aleatória e ARIMA.
* Especificamente projetado para lidar com dados climáticos do INMET, facilitando a análise e previsão de variáveis climáticas.

**Usuários primários:** O programa foi concebido principalmente para atender pesquisadores e profissionais especializados em ciência de dados e aprendizado de máquina, com foco em dados climáticos. Além disso, atende professores e estudantes das áreas de meteorologia, estatística, ciência de dados e engenharia de software que buscam uma ferramenta prática para avaliação de modelos com dados climáticos.

**Natureza do programa:** Este programa é uma prova de conceito parcial de uma ferramenta utilitária para a avaliação de modelos de aprendizado de máquina. Ele oferece uma base sólida para futuras expansões e refinamentos, podendo ser integrado a sistemas maiores ou utilizado de forma independente para fins educativos e de pesquisa.

**Ressalvas:**

* O programa assume que os dados de entrada estão limpos e devidamente formatados, e pode exigir ajustes manuais para casos específicos.
* O desempenho dos modelos de séries temporais como ARIMA pode variar significativamente dependendo da parametrização, sendo recomendada uma análise cuidadosa dos dados antes do uso.
* Algumas funcionalidades podem requerer bibliotecas específicas que não estão inclusas no escopo básico do programa, como statsmodels para ARIMA e scikit-learn para outros modelos.

Este programa fornece uma ferramenta robusta para a análise e avaliação de modelos de aprendizado de máquina, oferecendo suporte a uma variedade de cenários e necessidades de pesquisa e ensino, com um foco particular em variáveis climáticas fornecidas pelo INMET.