**O que é uma rede neural?**

Uma rede neural é um modelo que toma decisões de uma forma semelhante ao cérebro humano, utilizando processos que imitam a maneira como os neurônios biológicos trabalham juntos para identificar fenômenos, avaliar opções e chegar a conclusões.

Toda rede neural é composta por camadas de nós, ou neurônios artificiais: uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Cada nó se conecta a outros e possui seu próprio peso e limiar associados. Se a saída de qualquer nó individual estiver acima do valor de limiar especificado, esse nó é ativado, enviando dados para a próxima camada da rede. Caso contrário, nenhum dado é passado para a próxima camada.

[**Redes neurais recorrentes (RNNs)**](https://www.ibm.com/br-pt/topics/recurrent-neural-networks) são identificadas por seus loops de feedback. Esses algoritmos de aprendizado são principalmente utilizados ao lidar com dados de séries temporais para fazer previsões sobre resultados futuros, como previsões do mercado de ações ou previsões de vendas.

Arquiteturas RNN variantes

As variantes populares da arquitetura RNN incluem

* **Redes neurais recorrentes bidirecionais (BRRNs)**
* **Memória de curto longo prazo (LSTM)**
* **Unidades recorrentes fechadas (GNUs)**

**Redes LSTM**

Os LSTMs são explicitamente projetados para evitar o problema de dependência de longo prazo. Lembrar informações por longos períodos de tempo é praticamente o seu comportamento padrão, não é algo que eles lutam para aprender!

As redes **Long Short-Term Memory** (memória de curto e longo prazo), usualmente chamadas simplesmente de “LSTMs”, são um tipo especial de rede neural recorrente, pois são capazes de aprender conexões de longo prazo. Dessa maneira, elas têm um incrível poder de predição e funcionam muito bem em uma variada gama de problemas, sendo amplamente usadas atualmente.

A arquitetura de uma LSTM, faz todo sentido do ponto de vista biológico. Como já sabemos, há diversas analogias entre as redes neurais biológicas e as redes neurais artificiais. Quando olhamos para o modo como o cérebro humano funciona, descobrimos que a nossa memória pode ser dividida em dois tipos:

* memória de curto prazo, que atua no momento em que a informação está sendo adquirida, retém essa informação por alguns segundos e, então, a destina para ser guardada por períodos mais longos, ou a descarta.
* memória de longo prazo, que retém de forma definitiva a informação, permitindo sua recuperação ou evocação. Nela estão contidos todos os nossos dados autobiográficos e todo nosso conhecimento. Sua capacidade é praticamente ilimitada.

Intuitivamente, é assim que as LSTMs funcionam. Por uma célula desse tipo de rede, existem dois canais, que podem ser interpretados da mesma maneira como interpretamos a memória humana!

ADAM (Adaptive Moment Estimation) é um algoritmo de otimização usado em aplicações de aprendizado de máquina e aprendizado profundo. É uma combinação de duas metodologias de descida de gradiente: RMSProp (Root Mean Square Propagation) e Momentum.

**Principais características**

1. **Taxas de aprendizado adaptativas** : Adam ajusta dinamicamente as taxas de aprendizado para cada parâmetro com base no primeiro e segundo momentos dos gradientes.
2. **Correção de viés** : Adam realiza a correção de viés para abordar o problema de estimativas tendenciosas do primeiro e segundo momentos, particularmente durante os períodos de treinamento inicial.

Adam é um poderoso algoritmo de otimização amplamente usado em aprendizado profundo para treinamento de redes neurais, incluindo modelos LSTM. Seu mecanismo de taxa de aprendizado adaptável e correção de viés o tornam eficaz para uma variedade de tarefas de otimização, fornecendo convergência mais rápida e desempenho aprimorado em comparação aos métodos de otimização tradicionais. Ao usar Adam no código fornecido, o modelo classificador LSTM pode aprender eficientemente com os dados de treinamento e fazer previsões precisas para tarefas de classificação de sequência.

A classe sequencial do Keras, amplamente utilizada para construção de modelos simples. Porém, apesar de simples, esta classe consegue construir a maioria das redes neurais com que você provavelmente já trabalhou. É uma classe realmente fácil e intuitiva de se trabalhar, possibilitando construir modelos de Deep Learning para as mais variadas tarefas: classificação, regressão, entre outros.

Também é necessário criar uma a camada Densa, pois ela possibilitará construir modelos de Redes Neurais Profundas personalizáveis. A camada densa é responsável por conectar todos os neurônios de uma camada anterior a todos os neurônios de uma camada posterior. A camada densa contém parâmetros, que são usados para ajustar a saída de cada neurônio. Esses parâmetros são aprendidos durante o treinamento da rede neural. A camada densa também pode ser usada para adicionar não-linearidade ao modelo, o que é importante para modelos de redes neurais profundas.

**1. Erro Absoluto Médio (MAE): 0.3639**

**O que é o MAE?**

* O **MAE** mede a média dos valores absolutos das diferenças entre as previsões e os valores reais.
* Ele indica, em média, o quanto as previsões do modelo estão afastadas dos valores reais, sem considerar a direção dos erros (se são positivos ou negativos).

**Interpretação do Valor Obtido:**

* Um **MAE de 0.3639** significa que, em média, as previsões do seu modelo estão a aproximadamente **0,36 unidades monetárias** de distância dos valores reais.
* Como você está trabalhando com preços de fechamento em dólares, isso indica que, em média, o modelo erra por cerca de **36 centavos de dólar**.

**O que isso significa para o modelo?**

* Esse é um erro médio relativamente baixo, especialmente se comparado com o valor médio dos preços das ações no seu conjunto de dados.
* Indica que o modelo está fazendo previsões bastante próximas dos valores reais.

**2. Erro Quadrático Médio (MSE): 0.2002**

**O que é o MSE?**

* O **MSE** é a média dos quadrados dos erros entre as previsões e os valores reais.
* Ele penaliza erros maiores mais severamente devido ao quadrado dos erros, o que faz com que seja sensível a outliers.

**Interpretação do Valor Obtido:**

* Um **MSE de 0.2002** indica que a média dos quadrados dos erros é de aproximadamente **0,20**.
* Isso representa o erro quadrático médio em termos das unidades ao quadrado (neste caso, dólares quadrados).

**O que isso significa para o modelo?**

* Um MSE baixo sugere que os erros não são apenas pequenos em média, mas também que erros grandes (outliers) não estão influenciando significativamente o desempenho do modelo.
* Confirma que o modelo está fazendo previsões precisas e consistentes.

**3. Erro Quadrático Médio da Raiz (RMSE): 0.4474**

**O que é o RMSE?**

* O **RMSE** é a raiz quadrada do MSE e fornece uma medida do erro padrão das previsões.
* Está na mesma unidade dos dados originais (dólares), o que facilita a interpretação.

**Interpretação do Valor Obtido:**

* Um **RMSE de 0.4474** significa que o desvio padrão dos erros de previsão é de aproximadamente **$0,45**.
* Isso indica que, em média, as previsões do modelo estão a cerca de **45 centavos de dólar** de distância dos valores reais.

**Comparação com o Desvio Padrão dos Valores Reais:**

* O **desvio padrão dos valores reais** é **2.5034**, o que representa a variabilidade natural dos preços das ações em seu conjunto de dados.
* **Como o RMSE (0.4474) é significativamente menor que o desvio padrão dos valores reais (2.5034), isso indica que o modelo está prevendo os preços com uma precisão muito maior do que a variabilidade natural dos dados.**

**O que isso significa para o modelo?**

* O RMSE ser menor que o desvio padrão confirma que o modelo é eficaz em capturar os padrões nos dados e reduzir a incerteza nas previsões.
* Indica um bom ajuste do modelo, conforme mencionado na conclusão.

**4. Coeficiente de Determinação (R²): 0.9681**

**O que é o R²?**

* O **R²** mede a proporção da variância nos dados dependentes que é explicada pelo modelo.
* Varia entre 0 e 1, onde valores mais próximos de 1 indicam que o modelo explica melhor a variabilidade dos dados.

**Interpretação do Valor Obtido:**

* Um **R² de 0.9681** significa que o modelo explica **aproximadamente 96,81%** da variância nos preços das ações no conjunto de teste.
* Apenas cerca de 3,19% da variância nos dados não é explicada pelo modelo.

**O que isso significa para o modelo?**

* Um R² tão alto indica que o modelo tem um excelente poder preditivo e está capturando muito bem os padrões nos dados.
* Sugere que o modelo é confiável para prever os preços das ações dentro do contexto dos dados utilizados.

**5. Desvio Padrão dos Valores Reais: 2.5034**

**O que é o Desvio Padrão?**

* O **desvio padrão** é uma medida da dispersão ou variabilidade dos dados em relação à média.
* Indica o quanto os valores dos preços das ações variam em torno do valor médio.

**Interpretação do Valor Obtido:**

* Um desvio padrão de **2.5034** significa que os preços das ações geralmente variam em aproximadamente **$2,50** em torno da média.
* Isso representa a volatilidade natural dos preços no seu conjunto de dados.

**6. Conclusão Geral**

**O RMSE é menor que o desvio padrão, indicando um bom ajuste do modelo.**

**O que isso significa?**

* **RMSE < Desvio Padrão:** Quando o RMSE é menor que o desvio padrão dos valores reais, isso indica que o modelo está reduzindo significativamente a incerteza nas previsões em comparação com a variabilidade natural dos dados.
* **Bom Ajuste do Modelo:** O modelo está prevendo os preços com erros menores do que a variação típica dos preços, o que é um sinal de um modelo eficaz.

**Síntese dos Resultados**

* **Precisão nas Previsões:**
  + O MAE e o RMSE baixos indicam que o modelo está fazendo previsões muito próximas dos valores reais, com erros médios de aproximadamente **$0,36** e **$0,45**, respectivamente.
* **Explicação da Variância:**
  + O R² de **0.9681** mostra que o modelo explica quase toda a variabilidade nos preços das ações, o que é excelente.
* **Comparação com a Variabilidade dos Dados:**
  + O fato de o RMSE ser muito menor que o desvio padrão dos valores reais confirma que o modelo está superando a variabilidade natural dos preços, fornecendo previsões precisas.
* **Interpretação Global:**
  + **O modelo LSTM treinado está performando muito bem**, capturando os padrões nos dados e fornecendo previsões precisas dos preços das ações.
  + **Aplicabilidade Prática:** Esses resultados sugerem que o modelo pode ser confiável para aplicações práticas, como suporte à tomada de decisões de investimento, desde que sejam consideradas as limitações inerentes aos modelos preditivos e ao mercado financeiro.

Os resultados das métricas indicam que seu modelo LSTM está fazendo um excelente trabalho na previsão dos preços das ações, com erros médios baixos e alta capacidade de explicar a variância nos dados. Isso sugere que o modelo é uma ferramenta eficaz para prever movimentos de preços no contexto dos dados e período analisados.

**Resumo das Etapas Principais do Código**

Este projeto utiliza Redes Neurais Recorrentes do tipo LSTM para modelar e prever preços de ações com séries temporais. Aqui estão os **pontos principais e cruciais** do código:

**1. Importação de Dados**

* **Fonte dos Dados:** Utiliza a biblioteca yfinance para baixar os preços históricos da ação Petrobras (ticker: PBR) de 2004 a 2024.
* **Dados Obtidos:** Preços de abertura, fechamento, máxima, mínima, ajustado e volume negociado.
* **Objetivo:** Usar os preços de fechamento para modelagem.

**2. Análise Exploratória**

* **Verificação dos Dados:** Examina as primeiras linhas, a forma do dataset, verifica valores faltantes (.isna()), e obtém estatísticas descritivas (.describe()).
* **Correlação:** Analisa a correlação entre as variáveis.
* **Conclusão:** Não há valores faltantes, e a variável de interesse é Close.

**3. Pré-Processamento dos Dados**

* **Normalização:** Usa o MinMaxScaler para escalar os dados entre 0 e 1, o que é essencial para redes neurais.
* **Divisão dos Dados:**
  + **Treinamento:** 80% da série temporal.
  + **Teste:** 20% restante.
* **Separação em Entradas e Saídas:**
  + **Entradas (x\_train):** Sequências de 60 dias anteriores.
  + **Saídas (y\_train):** Preço do dia seguinte.

**4. Estrutura do Modelo**

* **Modelo LSTM:**
  + Duas camadas LSTM (50 neurônios cada).
  + Camada densa intermediária (25 neurônios).
  + Camada de saída com 1 neurônio (previsão).
* **Configuração:**
  + Otimizador adam e função de perda mean\_squared\_error.

**5. Treinamento do Modelo**

* **Treinamento:** Modelo é treinado com:
  + batch\_size=16
  + epochs=32
* **Resultado:** O modelo ajusta os pesos para minimizar o erro quadrático médio.

**6. Teste e Previsão**

* **Criação dos Dados de Teste:** São gerados conjuntos x\_test (entradas) e y\_test (valores reais).
* **Predições:** As previsões do modelo são escaladas de volta à escala original.

**7. Avaliação do Modelo**

* **Métricas Utilizadas:**
  + **Erro Absoluto Médio (MAE):** 0.364
  + **Erro Quadrático Médio (MSE):** 0.200
  + **Raiz do Erro Quadrático Médio (RMSE):** 0.447
  + **Coeficiente de Determinação (R²):** 96.8%
* **Conclusão:** O RMSE é menor que o desvio padrão dos valores reais (2.503), indicando um bom ajuste.

**8. Visualização dos Resultados**

* **Gráfico Comparativo:**
  + Mostra os preços de treinamento, validação e previsões no mesmo gráfico.
* **Médias:**
  + As médias dos valores previstos e reais são muito próximas.

**9. Previsões Futuras**

* **Janela Deslizante:**
  + Usa os últimos 5265 dias para prever 5 dias à frente.
* **Método:** Adiciona a previsão mais recente e remove o timestep mais antigo a cada iteração.
* **Resultado:**
  + Gera previsões para os 5 dias de dezembro de 2024.

**Conclusão**

Este projeto demonstra as etapas essenciais para modelar séries temporais com redes LSTM:

1. **Coleta e preparação dos dados.**
2. **Estruturação do modelo neural.**
3. **Treinamento e avaliação com métricas confiáveis.**
4. **Previsão futura usando janelas deslizantes.**

O modelo obteve **bom desempenho**, com erros baixos (RMSE: 0.447) e alta explicação da variância (R²: 96.8%). Pode ser usado como base para estudos mais avançados ou aplicações práticas em análise financeira.