LSTM 모델을 사용한 주식 예측

# 소개

이 문서는 LSTM(Long Short-Term Memory) 모델을 사용하여 주식 데이터를 학습하고 예측하는 전체 과정을 설명합니다. LSTM은 순차 데이터에서 중요한 패턴을 학습하는 데 효과적입니다. 이 문서에서는 데이터 전처리, 모델 학습, 결과 평가 및 저장에 대한 단계별 설명을 제공합니다.

# 데이터 준비

주식 데이터를 읽고 전처리하여 모델에 공급할 준비를 합니다. 데이터는 학습용, 검증용, 테스트용으로 나누어집니다. 이 단계에서는 데이터의 정규화 및 텐서 변환도 포함됩니다.

## 데이터 로드 및 전처리

1. 로컬 데이터 폴더에서 주식 데이터를 읽어옵니다.  
2. 데이터를 학습용, 검증용, 테스트용으로 나눕니다.  
3. 각 데이터셋을 Min-Max 스케일링을 사용하여 정규화합니다.  
4. 데이터를 LSTM 모델이 학습할 수 있는 형태로 변환합니다.

# 모델 설정

LSTM 모델을 정의하고, 손실 함수와 최적화 알고리즘을 설정합니다. 모델은 입력 크기, 은닉 상태 크기, 레이어 수 등의 하이퍼파라미터로 초기화됩니다.

## LSTM 모델 정의

LSTM 모델은 PyTorch의 nn.Module을 상속받아 정의됩니다. 모델은 LSTM 계층과 출력 계층(선형 계층)으로 구성됩니다.

# Forward 메서드

forward 메서드는 모델의 순전파 과정을 정의합니다. 입력 데이터를 받아 LSTM 계층을 거쳐 최종 출력을 반환합니다. 이 메서드는 학습, 검증, 테스트 과정에서 호출됩니다.

# 모델 학습 과정

1. 모델을 학습 모드로 설정하고, 각 배치에 대해 순전파와 역전파를 수행하여 가중치를 업데이트합니다.  
2. 에포크마다 학습 손실을 계산하고 출력합니다.  
3. 검증 데이터를 사용하여 모델의 성능을 평가합니다.  
4. 조기 종료를 적용하여 최적의 모델을 저장합니다.

# 모델 평가 및 예측

모델이 학습된 후, 테스트 데이터를 사용하여 예측을 수행합니다. 예측 결과를 실제 값과 비교하여 성능을 평가합니다.

# 결과 저장 및 시각화

학습과 검증 손실을 그래프로 시각화하여 저장합니다. 예측 결과를 CSV 파일로 저장하고, 스케일러와 전체 데이터를 피클 파일로 저장합니다.

# 결론

이 문서에서는 LSTM 모델을 사용하여 주식 데이터를 예측하는 전체 과정을 설명했습니다. 데이터 준비, 모델 학습, 결과 평가 및 저장을 통해 효과적인 예측 모델을 구축할 수 있습니다.