Predicting Nikkei 225 Next-Day Closing Price

– by Yifan Shen

Data Preparation

Data Source

- Japanese Stock Nikkei 225 Index
- Federal Reserve Economic Data API via pandas_datareader
- Interval: May 1 2020 – May 1 2025 (~1,200 trading days)

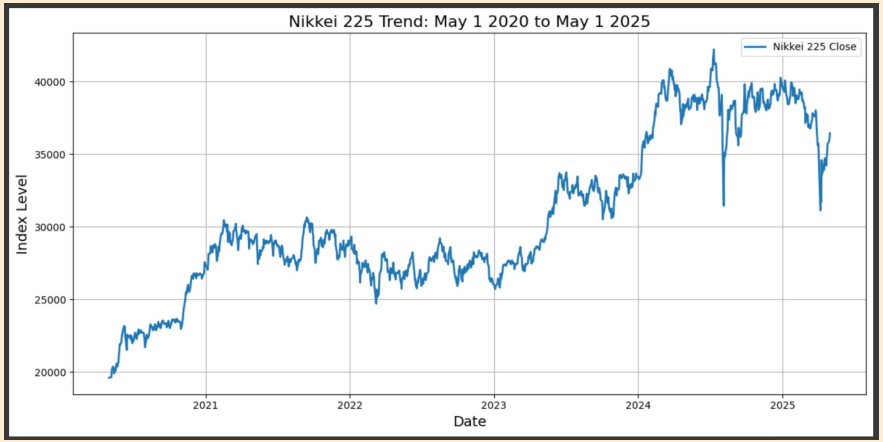
Preprocessing

- Reindex to all business days
- Linear interpolation to fill weekend/holiday gaps
- Drop any leading NaNs (before first real trading day)

	DATE	NIKKEI225
0	2020-05-01	19619.350
1	2020-05-04	19633.205
2	2020-05-05	19647.060
3	2020-05-06	19660.915
4	2020-05-07	19674.770
5	2020-05-08	20179.090
6	2020-05-11	20390.660
7	2020-05-12	20366.480
8	2020-05-13	20267.050
9	2020-05-14	19914.780
10	2020-05-15	20037.470
11	2020-05-18	20133.730
12	2020-05-19	20433.450
13	2020-05-20	20595.150
14	2020-05-21	20552.310
15	2020-05-22	20388.160
16	2020-05-25	20741.650
17	2020-05-26	21271.170
18	2020-05-27	21419.230
19	2020-05-28	21916.310

Exploratory Data Analysis

- Long-term rise: From ~19,500 to ~36,000 (May 2020–2025)
- Major dips: Early 2022 & late 2024 drawdowns with quick rebounds
- Volatility clusters: Spikes during crisis periods vs. calmer mid-2023
- Calendar effects: Small seasonal pullbacks (Jan/Feb, summer)
- Data continuity: Interpolated weekends/holidays, no gaps



Model Proposal

01 Objective

Use the last 10 days of Nikkei 225 closes to predict the next day's closing price

02 Model

User Input: 10-day sequences
Architecture: single LSTM layer → single dense output

03 Training Plan

80 % train / 20 % test split
Early stopping on validation MAE
Evaluate with MSE, MAE, and R^2

THANK YOU