Introduction to Stock Price Prediction

Big Picture: We Can Only Guess

In the stock market, nobody knows tomorrow's price for sure. The best we can do is make a **guess** based on what we have seen so far.

In this note we will:

- Look at a real stock (Apple).
- Try two very simple ways to guess the next day's price.
- Use lots of stories, like a child walking on a path.

The Data: AAPL Prices Over Time

Think of the stock price as a child walking along a road: each day the child stands at some position. We call the closing price on day t by the symbol P_t.

We can draw a picture of these prices over time: the horizontal axis is time (days), the vertical axis is price (in dollars).

We want to answer a very simple question:

If today the child stands at position P_t, where might the child be tomorrow?

1 Method 1: Naïve Forecast ("Tomorrow = Today")

Idea in Words

The simplest possible rule is almost silly:

Naïve rule: *"Tomorrow's price will be the same as today's price."*

If today's closing price is \$230, we simply guess that tomorrow will also be about \$230.

In symbols, if today's price is P_t, our guess for tomorrow (day $t+1$) is

$$\hat{P}_{t+1} = P_t.$$

Here $\hat{P}_{t+1}$ means "our predicted price for day $t+1$".

Child-on-a-Path Story

Imagine the price as a child walking along a path:

- Today the child is standing at position P_t.
- The naïve rule says: "Tomorrow the child will stand in *exactly the same spot*." We pretend the child does not move.

Of course real children move, and so do prices. But for *one small step* (a single day), this frozen-picture guess is surprisingly hard to beat in many markets.

Picture: Naïve Forecast vs Actual

Figure 1 shows the naïve forecast for the last 250 trading days of AAPL. The solid blue line is the **actual** price, and the dashed orange line is the **naïve forecast**.

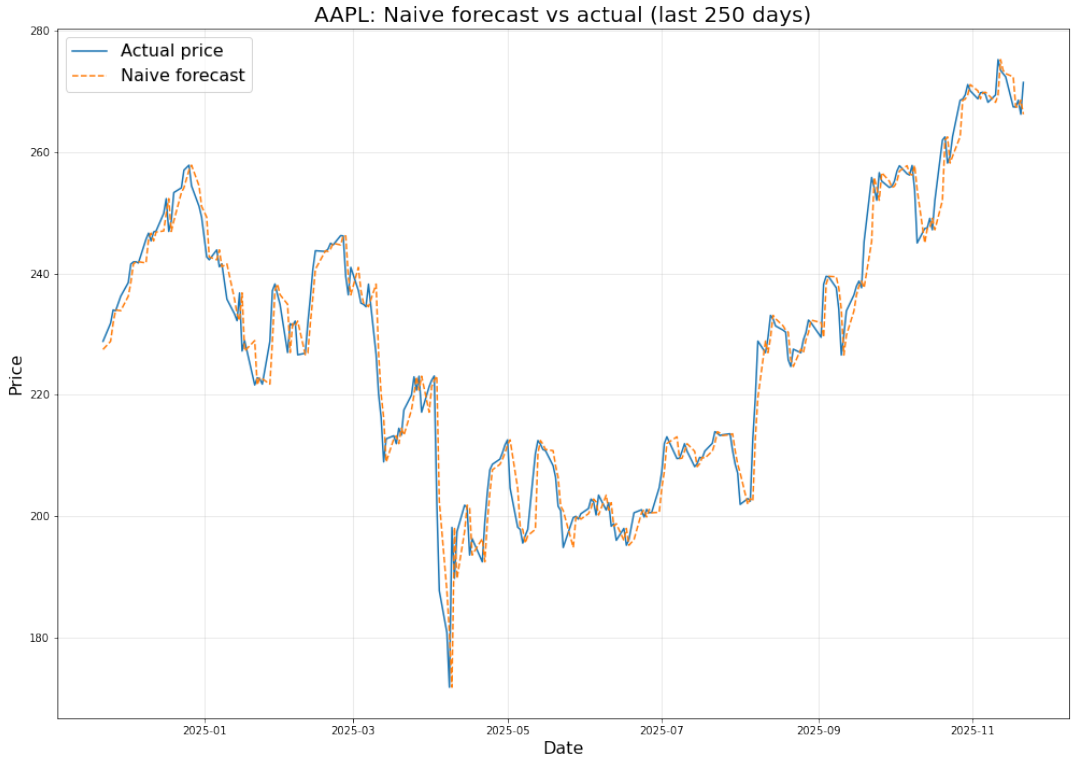

Figure 1: AAPL: Naïve forecast vs actual price (last 250 days).

What we see:

- The orange line hugs the blue line very closely.
- But it is always shifted by **one day**: it repeats *yesterday*'s price.
- When there is a sudden jump or drop, the naïve forecast is always "late" by one day.

So the naïve model is like a friend who only knows what happened yesterday and blindly repeats it.

2 Method 2: Moving Average Forecast ("Tomorrow = Recent Average")

Idea in Words

A small improvement (of course not always) is to look at the *last few days*, not just today.

> **Moving average rule:** *"Tomorrow's price will be close to the average of the last N days."*

For example, if we take the last 10 days, we compute their average and use that as tomorrow's price.

Formula, Explained Gently

Suppose we choose a window of N days. The closing prices for the last N days are

$$P_{t-N+1}, P_{t-N+2}, \ldots, P_{t-1}, P_t.$$

The **N-day simple moving average (SMA)** at time t is:

$$\text{SMA}_t = \frac{1}{N} \sum_{k=t-N+1}^{t} P_k.$$

This means:

- Take the last N prices.
- Add them up.
- Divide by N.

As a forecast, we say:

$$\hat{P}_{t+1} = \text{SMA}_t.$$

So tomorrow's price is guessed to be the "centre" of recent prices.

Child-on-a-Path Story

Again think of the child walking along a path.

- The naïve rule freezes the child: "Tomorrow they are in the same spot."
- The moving average rule looks at the child's **last N footprints** and says:

  > "These footprints are roughly around here. Tomorrow the child will probably be near this middle point."

So the moving average does not try to guess big jumps. It simply smooths out the noise and says "stay near the recent centre".

Picture: Moving Average Forecast vs Actual

Figure 2 shows the **10-day moving average forecast** for the last 250 trading days of AAPL. The dashed line is the forecast using a 10-day window.

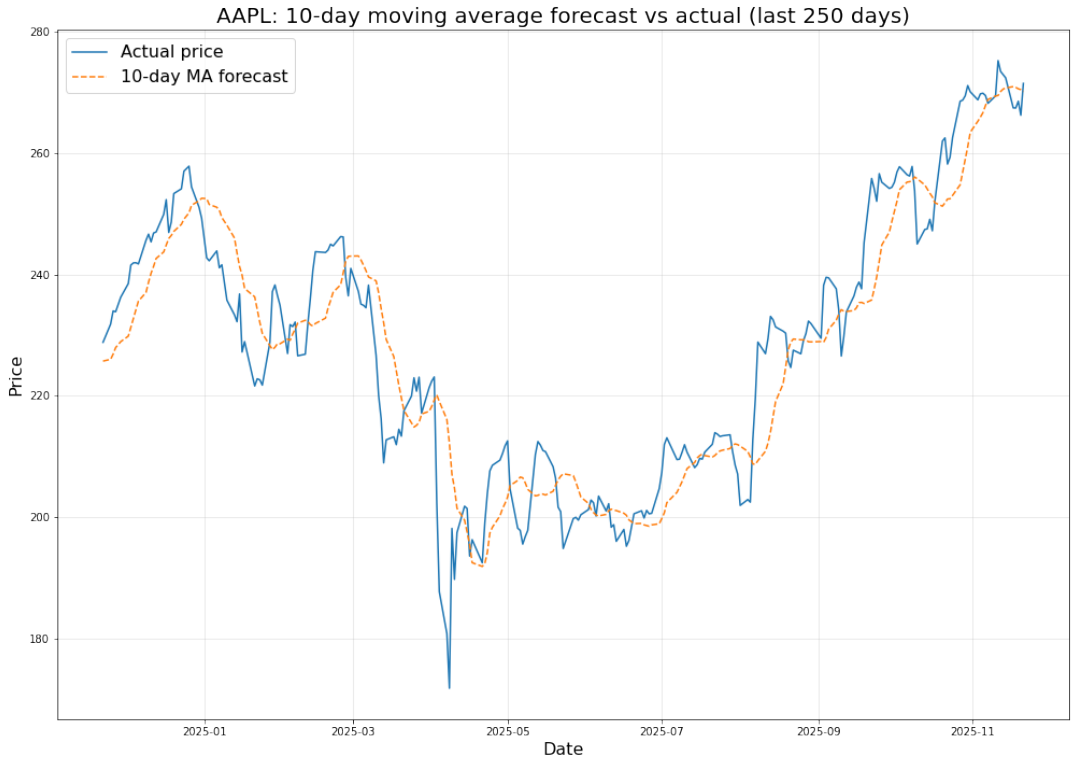

Figure 2: AAPL: 10-day moving average forecast vs actual price (last 250 days).

What we see:

- The moving average line is **smoother** than the actual price. It does not jump as wildly.
- When the price is trending up, the moving average follows, but more slowly.
- When the price suddenly falls, the moving average comes down more gently, because it remembers the earlier, higher prices.

So the moving average is like a slow, heavy train: once it is moving in some direction, it does not change direction very quickly.

3 Putting Both Methods Together

All Three Lines on One Plot

Figure 3 shows all three curves for the last 250 days:

- Blue: actual price.
- Orange dashed: naïve forecast.

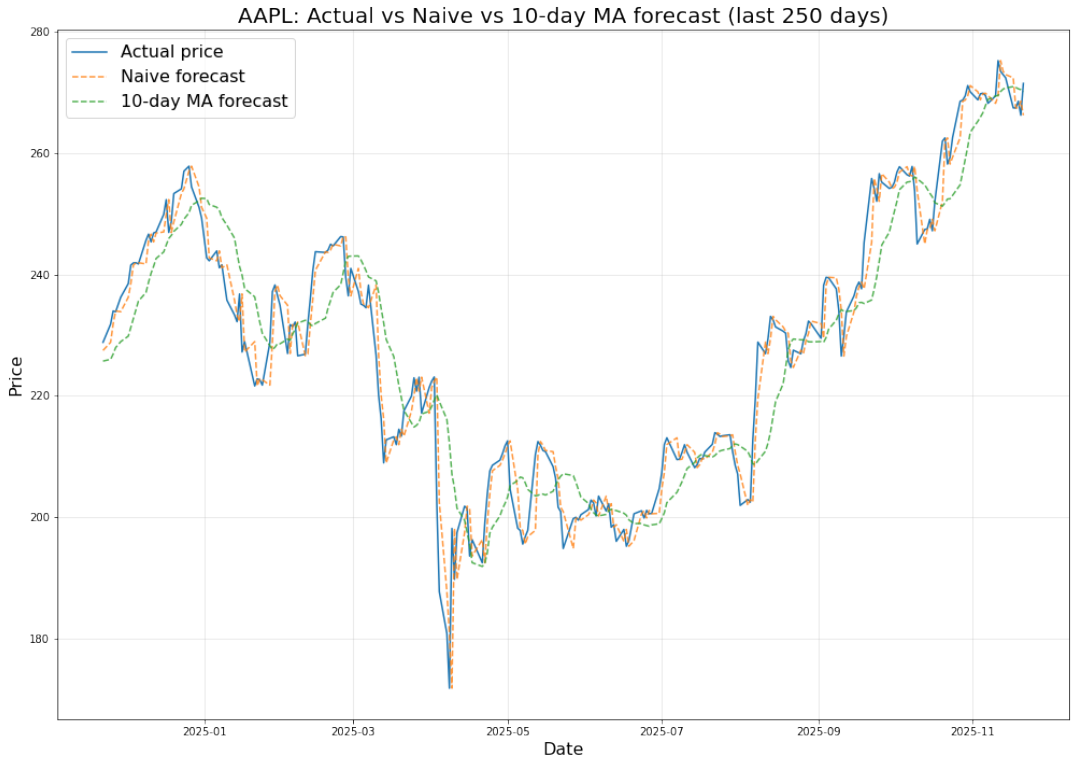

Figure 3: AAPL: Actual vs naïve vs 10-day moving average forecasts.

- Green dashed: 10-day moving average forecast.

Reading the picture:

- The naïve line (orange) hugs the blue line very tightly, but is always "one day late".
- The moving-average line (green) is smoother and reacts more slowly.
- During quiet periods, all three lines are close together.
- During fast moves (sharp up or down), both forecasts lag behind the real price.

How Wrong Are We, On Average?

To summarise how wrong a forecast is, we can use a simple number called the **mean absolute error** (MAE). You can think of it as:

> "On an average day, how many dollars are we off?"

If P_t is the actual price and $\hat{P}_t$ is our forecast on day t, then over N days:

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^{N} |P_t - \hat{P}_t|.$$

The computer can calculate this easily for both methods. Sometimes the moving average has smaller MAE (better), sometimes the naïve rule wins. There is no magic: real markets are noisy.

4 Why Start With These Simple Methods?

You might wonder: these rules are so simple, can they really be useful?

Yes, for a few reasons:

1. **They are clear.** You can explain them to a child: "Tomorrow = today" or "tomorrow = average of the last few days."
2. **They give a baseline.** Any fancy method should at least be better than these simple rules. If a complex model cannot beat the naïve forecast, it is not very helpful.
3. **They teach key ideas.**
   - using the **past** to guess the **future**;
   - smoothing noisy data with averages;
   - measuring **error** with something like MAE.

5 Beyond Naïve and Moving Averages

Real traders and researchers use many more tools. Here are a few, in simple language (no formulas needed here):

- **Exponential moving averages (EMA).** Like the simple moving average, but gives *more weight* to recent days and less to older days. This makes the forecast react faster.
- **Trend lines and regression.** Draw a straight line that best fits the past prices (a bit like drawing an average slope). Extend the line a little into the future.
- **Models on returns.** Instead of predicting the price itself, some models predict the *daily change* (return). For example, they may say: "If yesterday was strongly up, today is slightly more likely to be up."
- **Volatility models.** Some methods focus on predicting how *bumpy* the road will be (volatility), rather than the exact level of the price.
- **Machine learning and deep learning.** Fancy models (neural networks, trees, etc.) that try to find patterns in lots of data: prices, volumes, news, and more. These can be very powerful but also easy to over-trust or overfit.

Even when using these advanced methods, professionals still compare their results to simple baselines like the naïve forecast and moving averages.

Final Message

Predicting stock prices is like guessing where a child will be on a path:

- Very short term: the child is usually near where they are now.

- Recent average: the child often stays near the middle of their last steps.

In maths language, this becomes:

$$\hat{P}_{t+1} = P_t \quad \text{(naïve)}, \qquad \hat{P}_{t+1} = \text{average of last } N \text{ prices} \quad \text{(moving average)}.$$

These rules are not a way to get rich quickly. They are simple, honest starting points that help beginners see:

- how prediction is framed,
- how to check predictions against reality,
- and why even "dumb" baselines matter in real finance.

From here, you can slowly walk towards more advanced methods, always remembering that the future is uncertain, and every model is just a smarter guess.