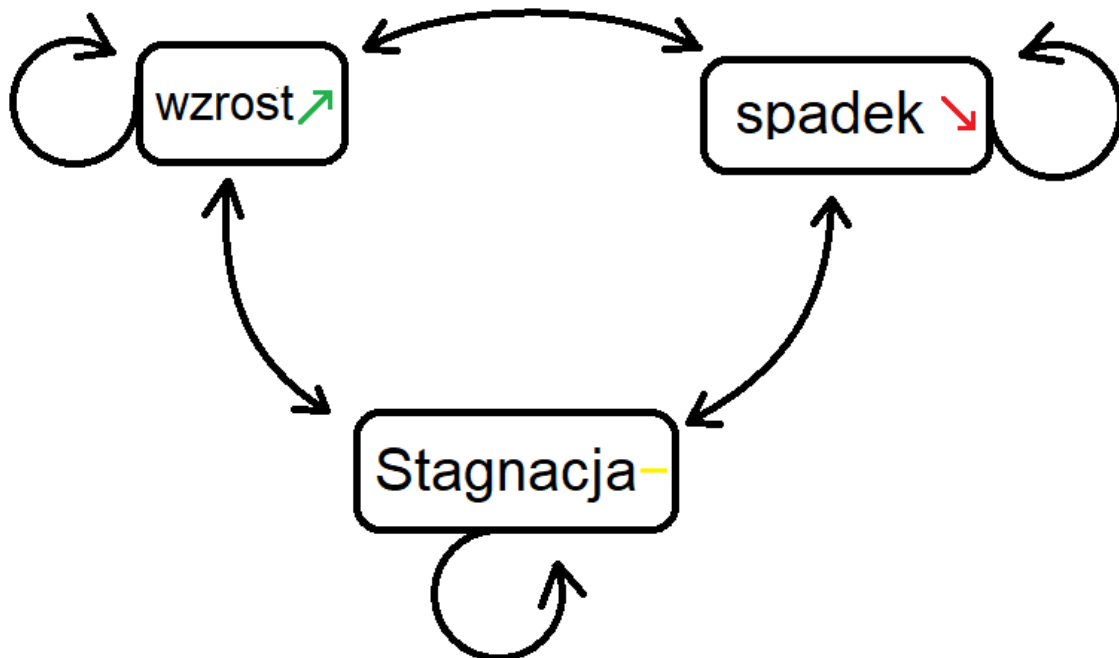


Symulacja akcji na giełdzie z wykorzystaniem łańcuchów Markowa

W pracy tej zostanie zbadana możliwość używania łańcuchów Markowa do symulacji przebiegu akcji na giełdzie.

Model



Model (macierz przejść) będzie opierał się na takim schemacie działania. Z danych historycznych, trzeba będzie więc obliczyć 9 prawdopodobieństw przejść; z danego stanu do każdego innego oraz do jego samego.

Dane

Skorzystano z danych dostępnych na stronie:

<https://www.kaggle.com/datasets/jacksoncrow/stock-market-dataset>

Zostały wybrane dane dotyczące jedynie portu Adani/Mundra, z lat 2007 - 2010. Brana pod uwagę była jedynie wartość akcji w trakcie zamknięcia giełdy w danym dniu. Ostatecznie tabela z danymi wyglądała następująco:

Date	Close
2007-11-27	962.90
2007-11-28	893.90
2007-11-29	884.20
2007-11-30	921.55
2007-12-03	969.30
...	...
2010-08-26	793.95
2010-08-27	768.40
2010-08-30	765.20
2010-08-31	768.65
2010-09-01	776.70

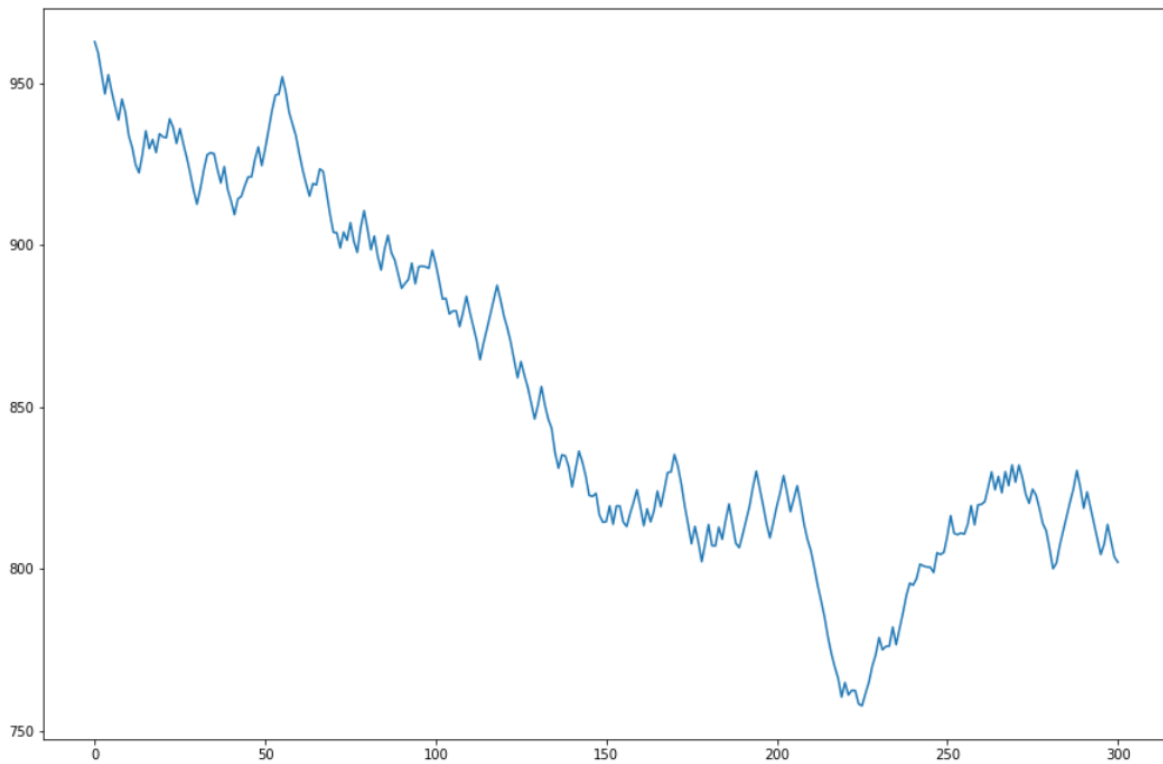
Obliczanie prawdopodobieństw — uczenie modelu

Zostały obliczone zmiany pomiędzy kolejnymi dniami i zostały one zaklasyfikowane jako stagnacja, wzrost lub spadek akcji. Stagnacja była klasyfikowana wtedy, gdy zmiana była w przedziale $<-2,2>$. Na podstawie przejść pomiędzy nimi zostały wyliczone prawdopodobieństwa do macierzy przejść. Ostatecznie macierz ta wyglądała następująco:

```
[0.48, 0.12, 0.40]
[0.40, 0.22, 0.38]
[0.36, 0.12, 0.52]
```

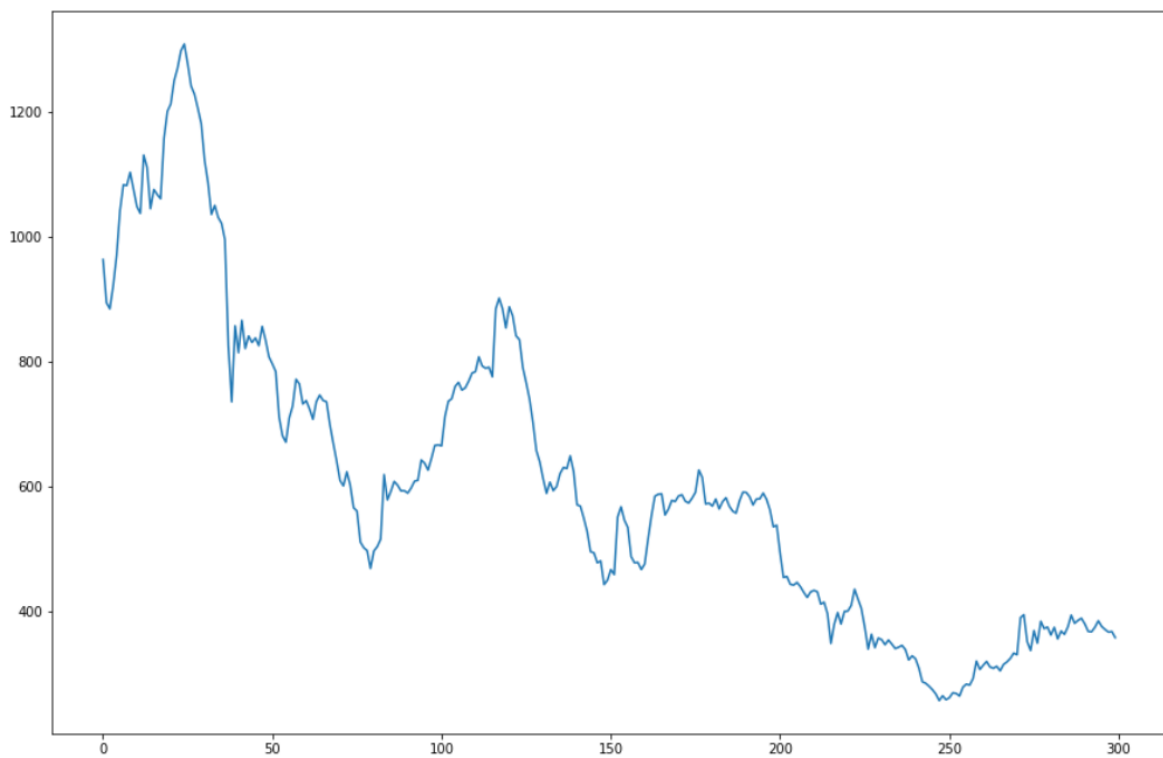
Gdzie wiersze i kolumny kolejno oznaczają: wzrost, stagnacja, spadek.

Symulacje



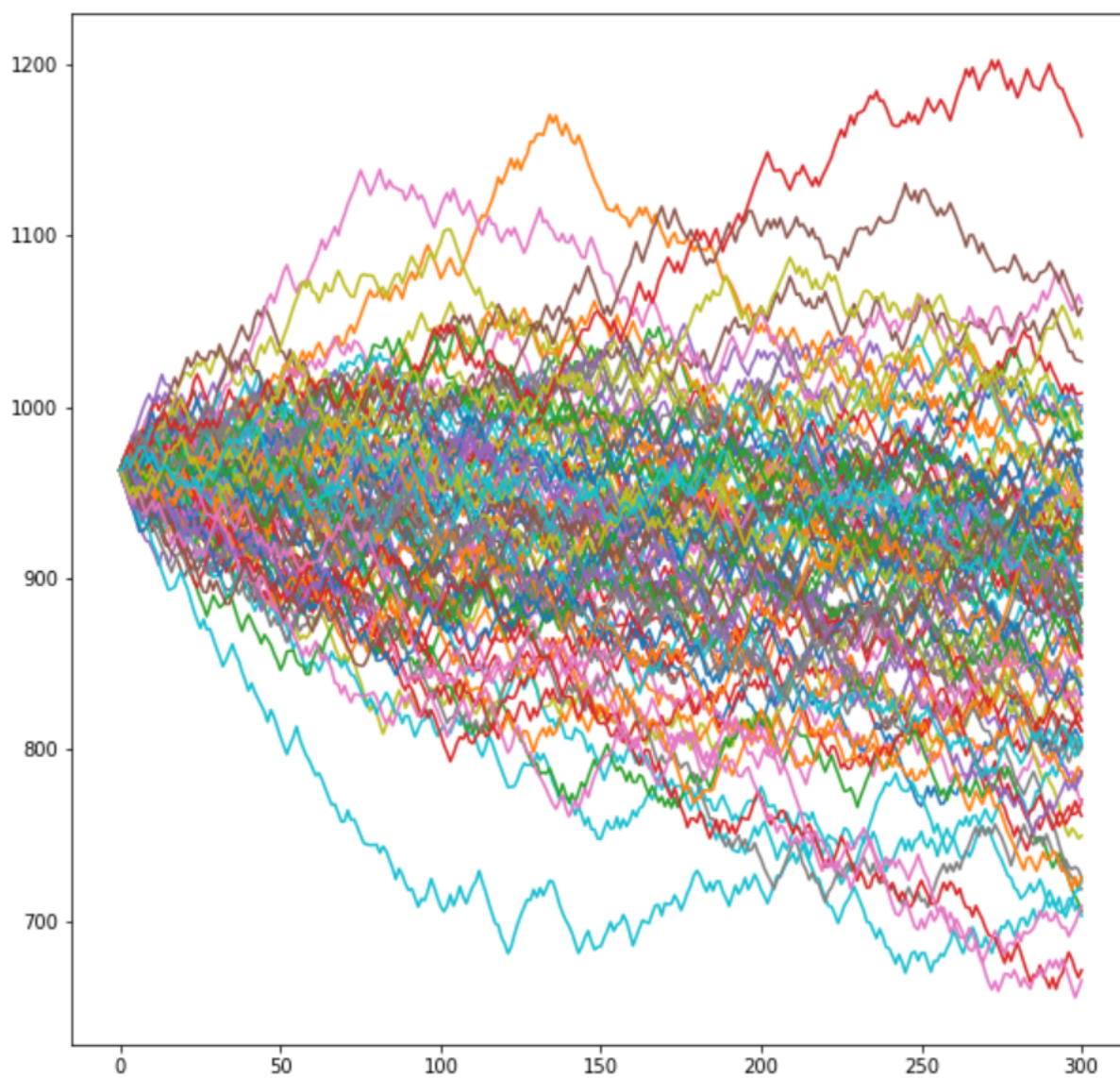
Przykładowa symulacja wykonana z wykorzystaniem powyższej macierzy przejść.

Dla porównania oryginalne wartości:



Jak model trochę poprawić — w rzeczywistości akcje spadły bardziej.

Można oczywiście wykonywać również wiele symulacji jednocześnie:



Wnioski

Największą zaletą wykonanej symulacji jest prawidłowe odwzorowanie trendu zawartego w danych. Patrząc jeszcze raz na poniższą macierz przejść, można zauważyć delikatną przewagę prawdopodobieństwa przejścia do stanu spadku wartości akcji, nad stanami stagnacji i wzrostem wartości akcji.

$$\begin{bmatrix} 0.48 & 0.12 & 0.40 \\ 0.40 & 0.22 & 0.38 \\ 0.36 & 0.12 & 0.52 \end{bmatrix}$$

Można to również zauważyć na wykresie wartości rzeczywistych, gdzie widać ewidentny spadek wartości na przestrzeni 300 dni. Symulacja uchwyciła tę zależność i mimo że skala i wartości spadku nie zostały odwzorowane w sposób idealny, to generalny trend danych został odwzorowany prawidłowo.

Wykres symulacji ma charakter dużo bardziej liniowy, niż wykres utworzony na podstawie oryginalnych danych. Symulacja utraciła dużo informacji na temat lokalnych procesów wewnątrz funkcji, takich jak wzrost wartości między 0 a 30 dniem, czy między 75 i 120 dniem. Symulacja zastąpiła te lokalne procesy wartościami zgodnymi z trendem globalnym funkcji. Można na podstawie tych informacji wywnioskować, że symulacje przy pomocy łańcuchów Markowa lepiej nadają się do symulowania ogólnego niż do odwzorowywania lokalnych procesów wewnątrz funkcji.

Posługując się wykresem wielu symulacji można wywnioskować kolejną istotną informację na temat stosowania łańcuchów markowa, a mianowicie istotność wykonywania wielu symulacji. Wartości poszczególnych symulacji bardzo różnią się od siebie, więc wykonując tylko jedną symulację, narażamy się na problemy związane z losowością wykonanej symulacji. Wykonana symulacja może nie odzwierciedlać w prawidłowy sposób danych, dlatego kluczowe jest wykonywanie wielu symulacji, a następnie przykładowo wybranie wartości średniej z wszystkich symulacji. Na ogół średnia z wielu symulacji będzie lepiej odzwierciedlała dane wejściowe niż pojedyncza symulacja.

Przykład zastosowania modelu w naukach o ziemi

Łańcuchy Markowa, znalazły zastosowanie w różnych dziedzinach nauki, nie tylko w analizie rynków finansowych. Okazują się one być niezmiernie przydatne w naukach o ziemi. Poniżej zostało wymienionych kilka przykładów użycia Łańcuchów Markowa dla tej właśnie dziedziny nauki.

- Analiza trendów klimatycznych
 - Modele oparte na łańcuchach Markowa mogą być użyteczne w analizie zmian klimatycznych na przestrzeni lat. Używając danych meteorologicznych, można stworzyć macierz przejść opisującą zmiany temperatury, opadów lub innych czynników wpływających na klimat. Symulacje na tej podstawie pozwoliłyby prognozować długoterminowe trendy klimatyczne oraz identyfikować prawdopodobieństwo przejścia między różnymi stanami atmosferycznymi.
- Prognozowanie Zagrożeń Naturalnych
 - Modele Łańcuchów Markowa mogą także być zastosowane do prognozowania zagrożeń naturalnych, takich jak trzęsienia ziemi czy erupcje wulkanów. Analiza prawdopodobieństw przejścia między różnymi stanami sejsmicznymi pozwala na identyfikację obszarów o podwyższonym ryzyku oraz lepsze przygotowanie do ewentualnych katastrof.
- Modelowanie Migracji Zwierząt
 - Model oparty na łańcuchach Markowa może być używany do analizy i prognozowania migracji różnych gatunków zwierząt. Przejścia między obszarami geograficznymi, klimatycznymi czy dostępnością pożywienia mogą być modelowane, co umożliwi lepsze zrozumienie zachowań migracyjnych i wpływu zmian środowiskowych na te procesy.

Przykładowy model łańcucha Markowa dla zmian pogody może obejmować trzy stany: słonecznie, deszczowo i pochmurnie. Załóżmy, że chcemy zamodelować zmiany pogody w kolejnych dniach. Poniżej przedstawiamy przykładowy prosty model łańcucha Markowa związany z tym scenariuszem:

- 1) Definicja stanów:
 - a) Słonecznie
 - b) Deszczowo
 - c) Pochmurnie
- 2) Macierz prawdopodobieństw przejść:
Przykładowa macierz przejścia P:

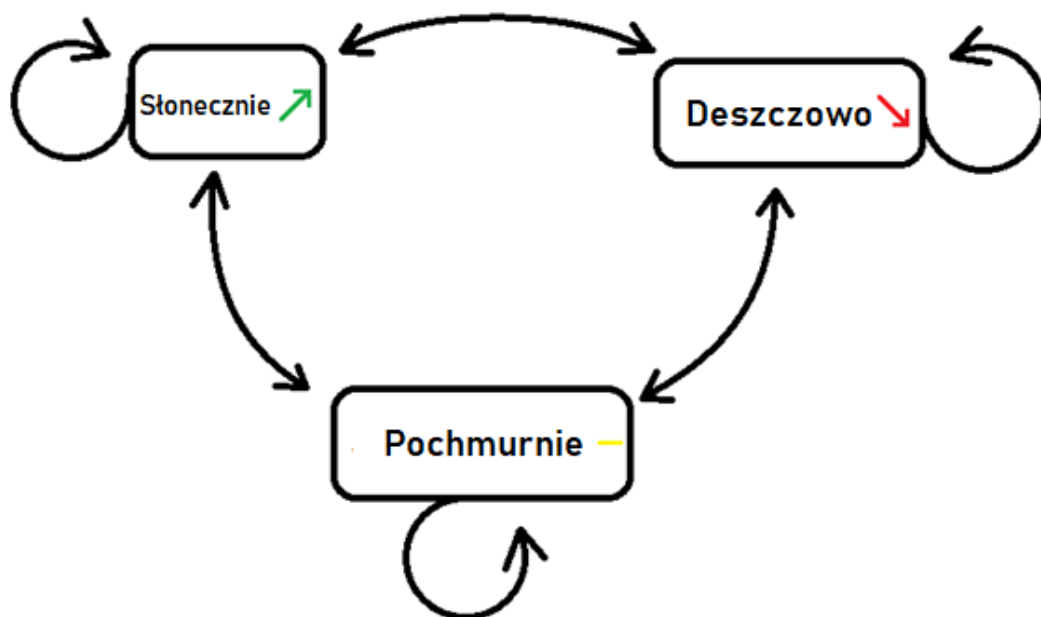
$$P = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}$$

3) Zastosowanie:

Przykładowe zastosowanie takiego modelu może obejmować prognozowanie pogody na podstawie aktualnego stanu. Na przykład, jeśli dzisiaj jest słonecznie, możemy użyć modelu, aby prognozować, że z 70% prawdopodobieństwem jutro również będzie słonecznie.

4) Obliczanie prawdopodobieństw dla historycznych danych:

Jeśli mamy dostęp do historycznych danych pogodowych, możemy użyć ich do obliczenia prawdopodobieństw przejść w macierzy P . Przykładowo, możemy policzyć, ile razy po słonecznym dniu nastąpił kolejny dzień ze słońcem, ile razy pojawił się deszcz, itp. Następnie normalizujemy te liczby do postaci prawdopodobieństw.



Przykładowy model.