

# Dokumentacja techniczna

## projektu z przedmiotu:

# „Architektury rozwiązań i wdrożeń SI”

### Opis projektu i jego funkcjonalności

W ramach tego projektu tworzony jest model AI. Dzięki temu modelowi będzie można przewidzieć czy następnego dnia będzie padał deszcz w Australii. Do nauki modelu zostały wykorzystane dane pomiarowe z platformy Kaggle (<https://www.kaggle.com/datasets/jsphyg/weather-dataset-rattle-package>). Jest to duża baza danych posiadająca dokładne dane pomiarowe z miejscowości rozmieszczonych po całej Australii. Zbiór danych wynosi około 145 tys rekordów. Do przygotowania modelu wykorzystaliśmy podane po niżej parametry:

1. Date:  
Data pomiaru warunków pogodowych w formacie RRRR-MM-DD.
2. Location:  
Nazwa konkretnej lokalizacji, dla której zebrane są dane pogodowe (np. "Albury").
3. MinTemp:  
Minimalna temperatura zarejestrowana w ciągu dnia (w stopniach Celsjusza). Na przykład, 13.4 oznacza, że najniższa temperatura tego dnia wynosiła 13.4°C.
4. MaxTemp:  
Maksymalna temperatura zarejestrowana w ciągu dnia (w stopniach Celsjusza). Na przykład, 22.9 oznacza, że najwyższa temperatura tego dnia wynosiła 22.9°C.
5. Rainfall:  
Suma opadów deszczu zarejestrowanych w ciągu dnia (w milimetrach). Na przykład, 0.6 oznacza, że tego dnia spadło 0.6 mm deszczu.
6. Evaporation:  
Całkowita ilość parowania klasy A zarejestrowana w ciągu dnia (w milimetrach). Dane mogą być niedostępne (NA), gdy brak jest odpowiednich pomiarów lub informacji.
7. Sunshine:  
Całkowita liczba godzin nasłonecznienia zarejestrowana w ciągu dnia. Podobnie jak w przypadku Evaporation, dane mogą być niedostępne (NA), gdy brak jest odpowiednich pomiarów lub informacji.
8. WindGustDir:  
Kierunek wiatru podczas największego porywu wiatru zarejestrowanego tego dnia (na przykład "W" oznacza Zachód).
9. WindGustSpeed:  
Prędkość wiatru podczas największego porywu zarejestrowanego tego dnia (w kilometrach na godzinę). Na przykład, 44 oznacza, że maksymalny poryw wiatru wyniósł 44 km/h.
10. WindDir9am:  
Kierunek wiatru o godzinie 9:00 rano (na przykład "W" oznacza Zachód).
11. WindDir3pm:  
Kierunek wiatru o godzinie 15:00 po południu (na przykład "WNW" oznacza Zachód-Północny Zachód).
12. WindSpeed9am:  
Prędkość wiatru o godzinie 9:00 rano (w kilometrach na godzinę). Na przykład, 20 oznacza, że wiatr wiejący o 9:00 rano osiągnął prędkość 20 km/h.

13. WindSpeed3pm:  
Prędkość wiatru o godzinie 15:00 po południu (w kilometrach na godzinę). Na przykład, 24 oznacza, że wiatr wiejący o 15:00 osiągnął prędkość 24 km/h.
14. Humidity9am:  
Wilgotność o godzinie 9:00 rano (w procentach). Na przykład, 71 oznacza, że o godzinie 9:00 rano wilgotność wynosiła 71%.
15. Humidity3pm:  
Wilgotność o godzinie 15:00 po południu (w procentach). Na przykład, 22 oznacza, że o godzinie 15:00 po południu wilgotność wynosiła 22%.
16. Pressure9am:  
Ciśnienie atmosferyczne o godzinie 9:00 rano (w hektopaskalach, hPa). Na przykład, 1007.7 oznacza, że ciśnienie o 9:00 rano wynosiło 1007.7 hPa.
17. Pressure3pm:  
Ciśnienie atmosferyczne o godzinie 15:00 po południu (w hektopaskalach, hPa). Na przykład, 1007.1 oznacza, że ciśnienie o 15:00 po południu wynosiło 1007.1 hPa.
18. Cloud9am:  
Zasłonięcie chmur o godzinie 9:00 rano (w oktach, gdzie 0 to brak chmur, a 8 to całkowite zasłonięcie). Na przykład, 8 oznacza, że o 9:00 rano było całkowicie pochmurno.
19. Cloud3pm:  
Zasłonięcie chmur o godzinie 15:00 po południu (w oktach). Na przykład, 8 oznacza, że o 15:00 po południu było całkowicie pochmurno.
20. Temp9am:  
Temperatura o godzinie 9:00 rano (w stopniach Celsjusza). Na przykład, 16.9 oznacza, że o 9:00 rano temperatura wynosiła 16.9°C.
21. Temp3pm:  
Temperatura o godzinie 15:00 po południu (w stopniach Celsjusza). Na przykład, 21.8 oznacza, że o 15:00 po południu temperatura wynosiła 21.8°C.
22. RainToday:  
Informacja czy wystąpiły opady deszczu tego dnia (No - Nie: 0, Yes - Tak: 1).
23. RainTomorrow:  
Informacja czy przewiduje się opady deszczu następnego dnia (No - Nie: 0, Yes - Tak: 1).

Dane te są dzielone na treningowe i testowe (80% treningowe, 20% testowe). Do wyboru najlepszego modelu użyliśmy AutoGluon. Jest to narzędzie, które trenuje różne modele na podstawie tych samych danych i wybiera najlepszy z nich. Dla wygody użytkownika użyliśmy narzędzie docker. Wystarczy jedno polecenie aby całe środowisko skonfigurowało się samodzielnie. W naszej aplikacji wykorzystaliśmy jeszcze Streamlit. Streamlit jest prostym narzędziem do tworzenia UI. Dzięki niemu użytkownik będzie mógł sprawdzić predykcje danego modelu oraz zobaczyć dane, które są do wykorzystywane do nauki jak i trenowania modeli.

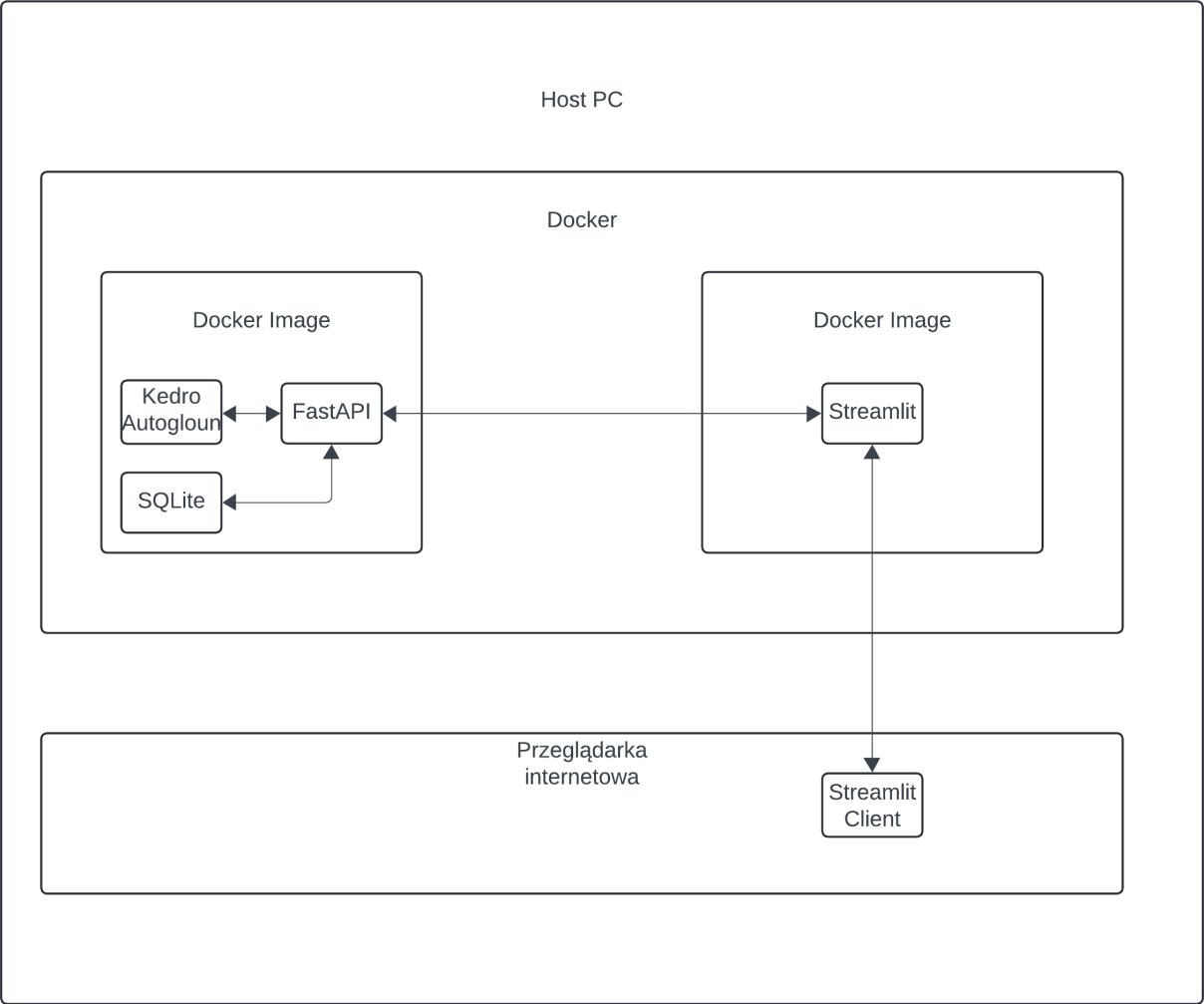
## Wykorzystane technologie:

- Kedro aby zarządzać pipeline'ami oraz aby przygotować dane
- SQLite służy do przechowywania danych
- AutoGluon uczenie i wybór najlepszego modelu
- WandDB śledzenie metryk trenowania
- FastAPI api do interakcji z modelami oraz bazą danych
- Streamlit do stworzenia prostego UI

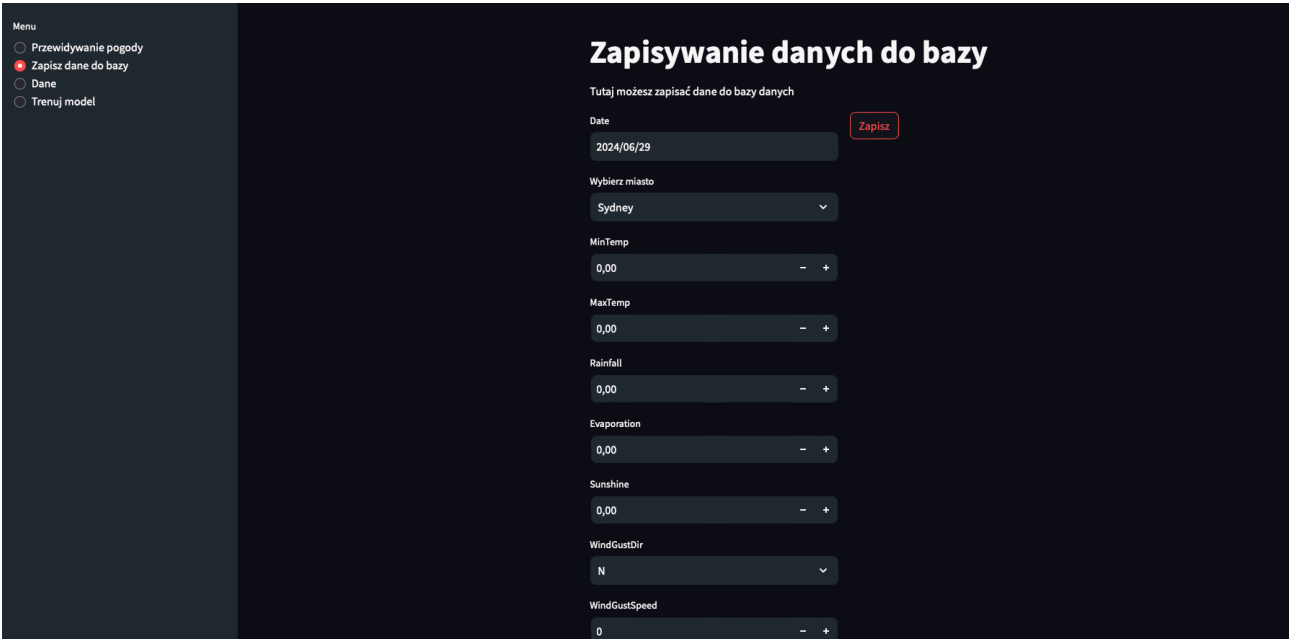
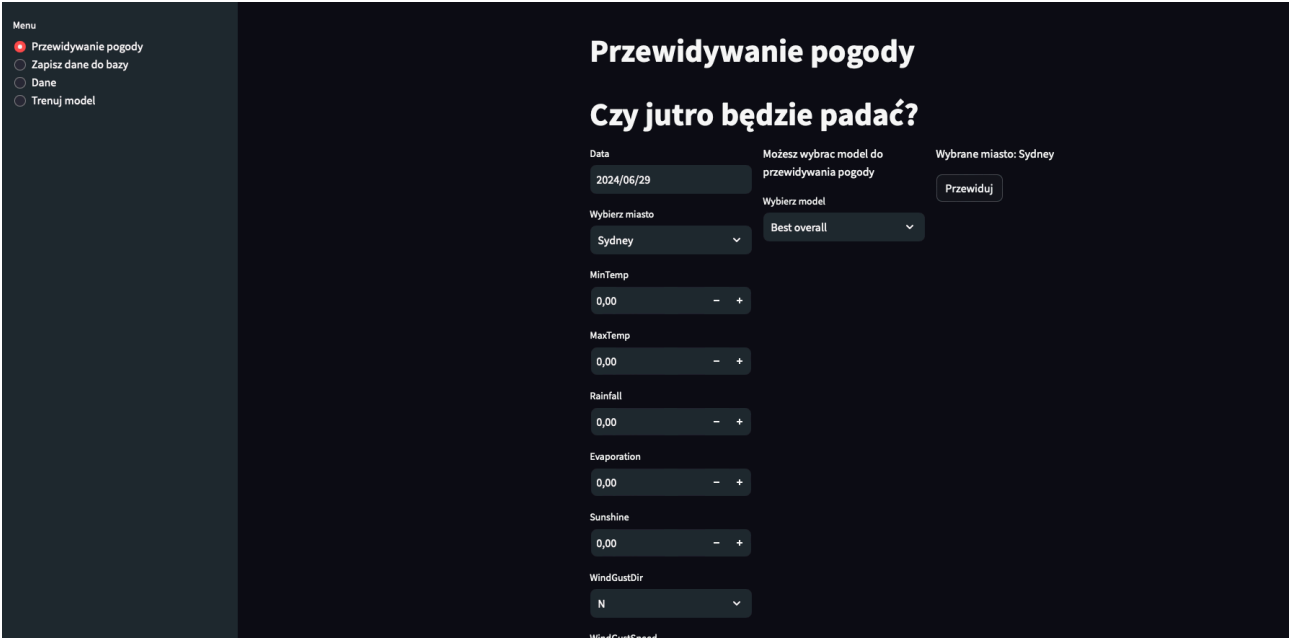
## Uruchomienie środowiska:

- uruchomienie docker
- ```
docker-compose up --build
```
- uruchomienie streamlit
- <http://localhost:8501/>

Architektura:



Zrzuty ekranu Streamlit:



Menu

Przewidywanie pogody

Zapisz dane do bazy

Dane

Trenuj model

Dane

Tutaj możesz zobaczyć dane

|   | date       | selected_city | min_temp | max_temp | rainfall | evaporation | sunshine | wind_gust_dir |
|---|------------|---------------|----------|----------|----------|-------------|----------|---------------|
| 0 | 2008-12-01 | Albury        | 13.4     | 22.9     | 0.6      | None        | None     | W             |
| 1 | 2008-12-02 | Albury        | 7.4      | 25.1     | 0        | None        | None     | WNW           |
| 2 | 2008-12-03 | Albury        | 12.9     | 25.7     | 0        | None        | None     | WSW           |
| 3 | 2008-12-04 | Albury        | 9.2      | 28       | 0        | None        | None     | NE            |
| 4 | 2008-12-05 | Albury        | 17.5     | 32.3     | 1        | None        | None     | W             |
| 5 | 2008-12-06 | Albury        | 14.6     | 29.7     | 0.2      | None        | None     | WNW           |
| 6 | 2008-12-07 | Albury        | 14.3     | 25       | 0        | None        | None     | W             |
| 7 | 2008-12-08 | Albury        | 7.7      | 26.7     | 0        | None        | None     | W             |
| 8 | 2008-12-09 | Albury        | 9.7      | 31.9     | 0        | None        | None     | NNW           |
| 9 | 2008-12-10 | Albury        | 13.1     | 30.1     | 1.4      | None        | None     | W             |