
Forecasting piogge - Emilia Romagna

Operational Analytics
2024/2025
G.Casamenti

Introduzione e obiettivi

La previsione delle piogge sono un problema critico per numerosi settori:

- Agricoltura e pianificazione delle irrigazioni,
- Gestione del rischio idrologico (alluvioni, siccità),
- Pianificazione urbana e ambientale.

Da questa motivazione nasce l'esigenza di studiare e confrontare diversi modelli di forecasting, tra cui

- Modelli statistici
- Reti neurali
- Alberi decisionali



L'obiettivo è determinare quale metodo offra il miglior compromesso tra accuratezza, stabilità e capacità di apprendere dinamiche stagionali e fluttuazioni settimanali.

Data e preprocessing - pt1

I dati storici utilizzati per l'addestramento dei modelli sono stati ottenuti da fonti meteorologiche ufficiali e certificate. Alcune delle fonti di riferimento includono:

- ARPAE
- ISPRA - Istituto Superiore per la Protezione e la Ricerca Ambientale

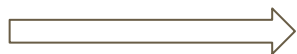
I dati sono stati aggregati su base settimanale e pre-processati per garantire coerenza temporale e qualità statistica.

Frequenza: settimanale

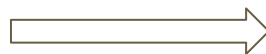
Unità di misura: mm di pioggia

Intervallo temporale: dati storici dal 2014 fino al 2023, previsione per tutto il 2024

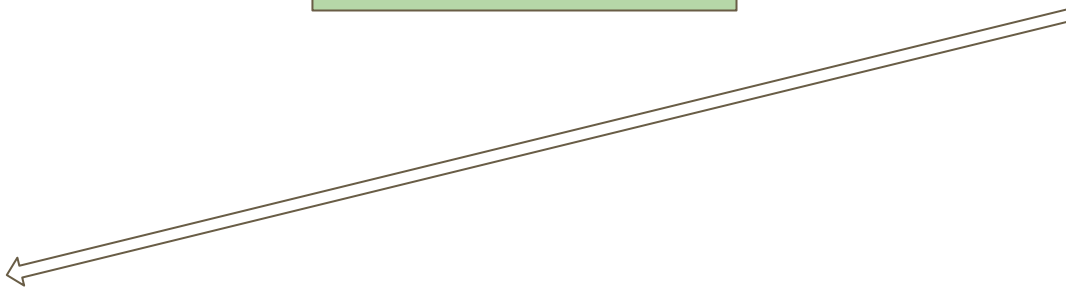
Data e preprocessing - pt2



Serie temporale
train e test



Analisi esplorativa dei
dati



Gestione outliers e
valori zero

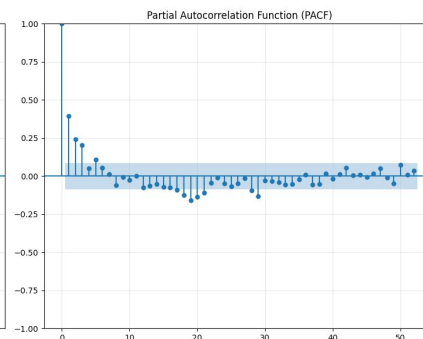
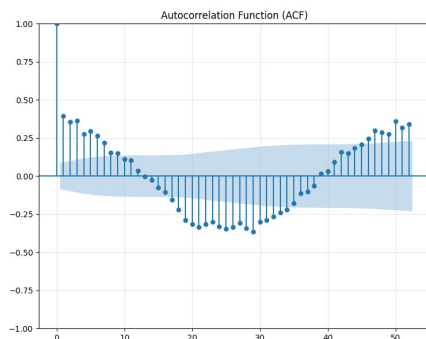
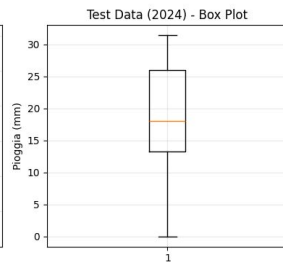
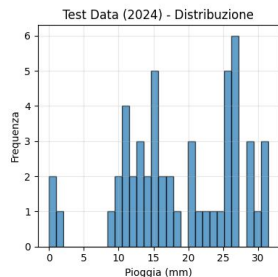
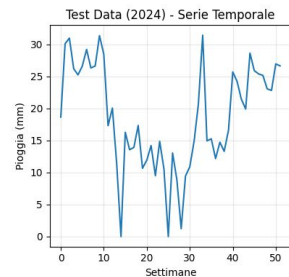
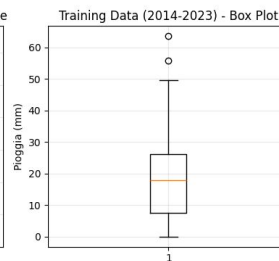
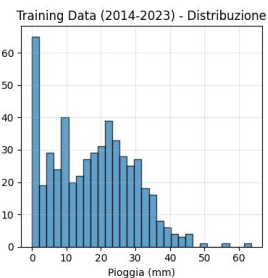
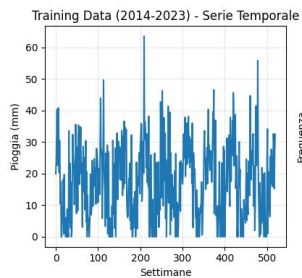
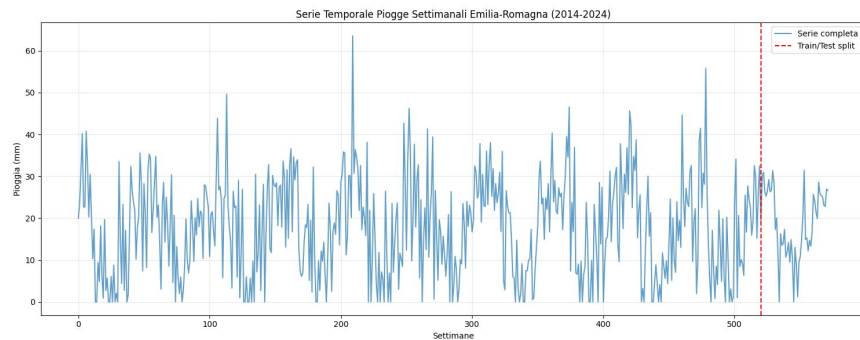


Test di stazionarietà



Analisi
autocorrelazione

Data e preprocessing - pt3



Modelli statistico - Sarima

- **Componenti non stagionali:** ARIMA(1,1,2)
- **Componenti stagionali:** ARIMA(1,1,1)
- **Stagionalità:** 52

--- RISULTATI ACCURATEZZA SARIMA ---

MAPE (Mean Absolute Percentage Error): 0.0955 (9.55%)

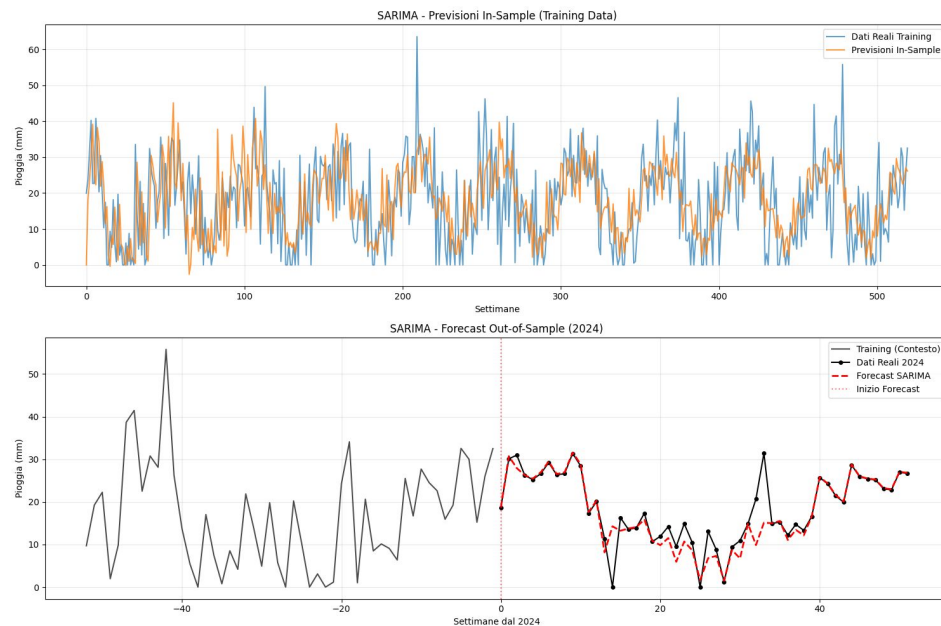
ME (Mean Error): -0.8993

MAE (Mean Absolute Error): 1.7397

MPE (Mean Percentage Error): -0.0756 (-7.56%)

RMSE (Root Mean Square Error): 3.7483

Correlazione: 0.9003



Rete neurale feedforward

Struttura:

- `nn.Linear(input_size, 64)`
- `nn.ReLU()`,
- `nn.Dropout(0.2)`,
- `nn.Linear(64, 32)`,
- `nn.ReLU()`,
- `nn.Linear(32, 1)`

--- RISULTATI ACCURATEZZA RETE NEURALE ---

MAPE (Mean Absolute Percentage Error): 0.2879 (28.79%)

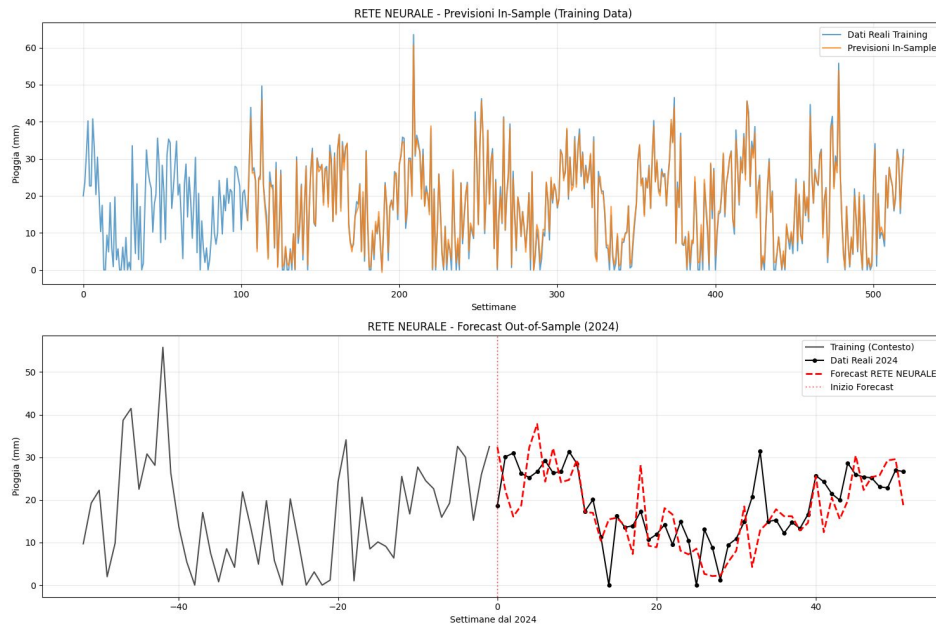
ME (Mean Error): -1.0646

MAE (Mean Absolute Error): 5.4412

MPE (Mean Percentage Error): -0.0602 (-6.02%)

RMSE (Root Mean Square Error): 7.1422

Correlazione: 0.6499

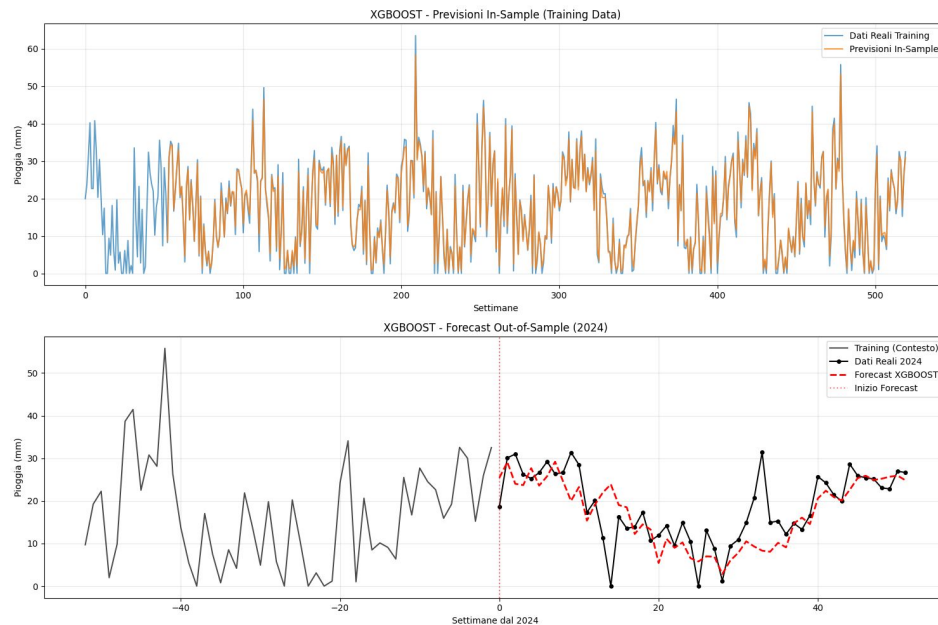


Decision tree - XGBoost

XGBRegressor:

- **objective='reg:squarederror',**
- **n_estimators=1000,**
- **max_depth=5,**
- **learning_rate=0.01,**
- **subsample=0.6,**
- **colsample_bytree=0.8,gamma=1**

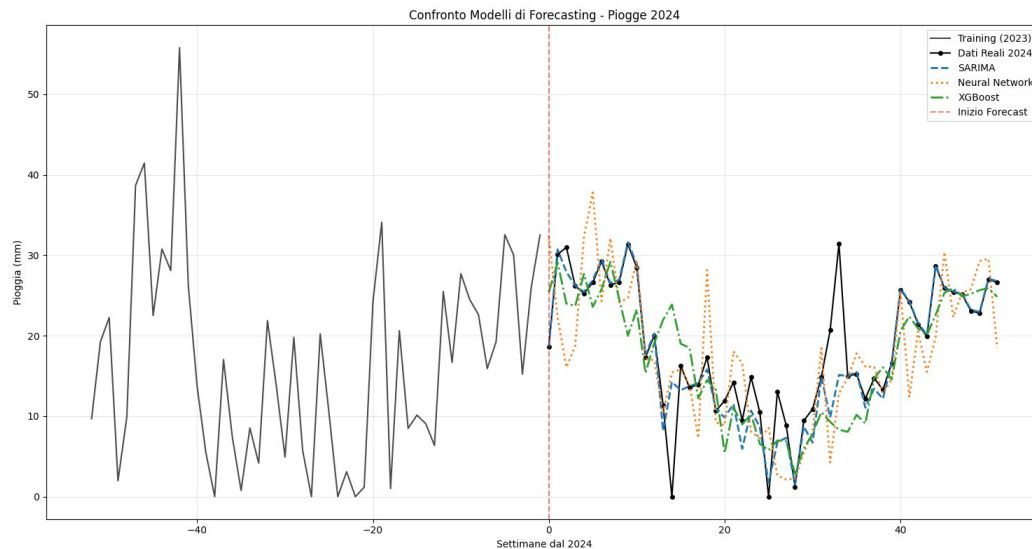
```
--- RISULTATI ACCURATEZZA XGBOOST ---  
MAPE (Mean Absolute Percentage Error): 0.2387 (23.87%)  
ME (Mean Error): -1.4770  
MAE (Mean Absolute Error): 4.2696  
MPE (Mean Percentage Error): -0.0770 (-7.70%)  
RMSE (Root Mean Square Error): 6.3283  
Correlazione: 0.6951
```



Conclusioni

Benchmark di valutazione:

Categoria	MAPE
ECCELLENTE:	5-15%
MOLTO BUONO:	15-25%
BUONO:	25-35%
ACCETTABILE:	35-50%
SCARSO:	50%+



Diebol-Mariano TEST

SARIMA vs Neural Network - DM stat: 5.2411, p-value: 0.0000

SARIMA vs XGBoost - DM stat: 2.9026, p-value: 0.0055

Neural Network vs XGBoost - DM stat: -1.0153, p-value: 0.3148