Application of KPI forecasting methods for engineering recruitment

Monia Houichi, Makrem Jannadi, Khaled Bendriss and Dr. Ines Abdeljaoued-Tej

WEVIOO,Tunisia

Higher School of Statistics and Information Analysis, University of Carthage, Tunisia

Laboratory BIMS, Institut Pasteur de Tunis, University of Tunis El Manar, Tunisia
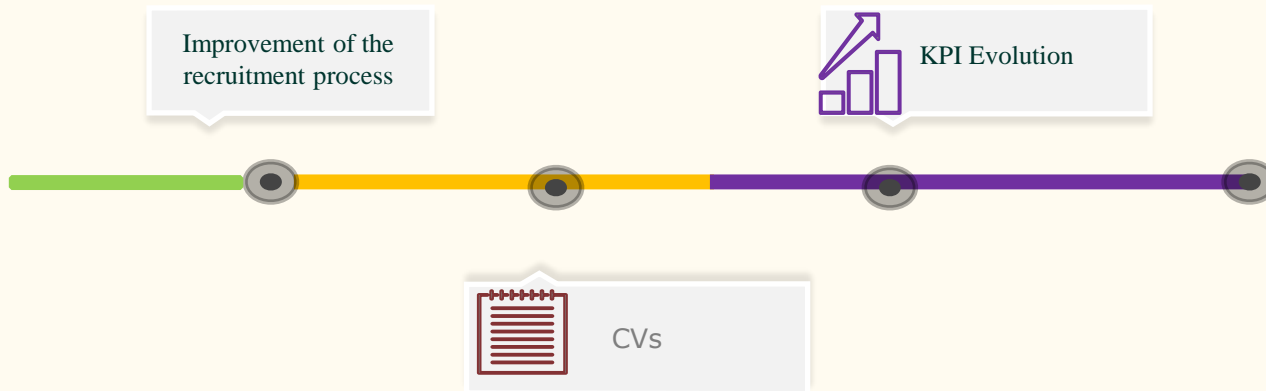
National Engineering School of Tunis, University of Tunis El Manar,Tunisia

Tunisia, september

Content

Introduction

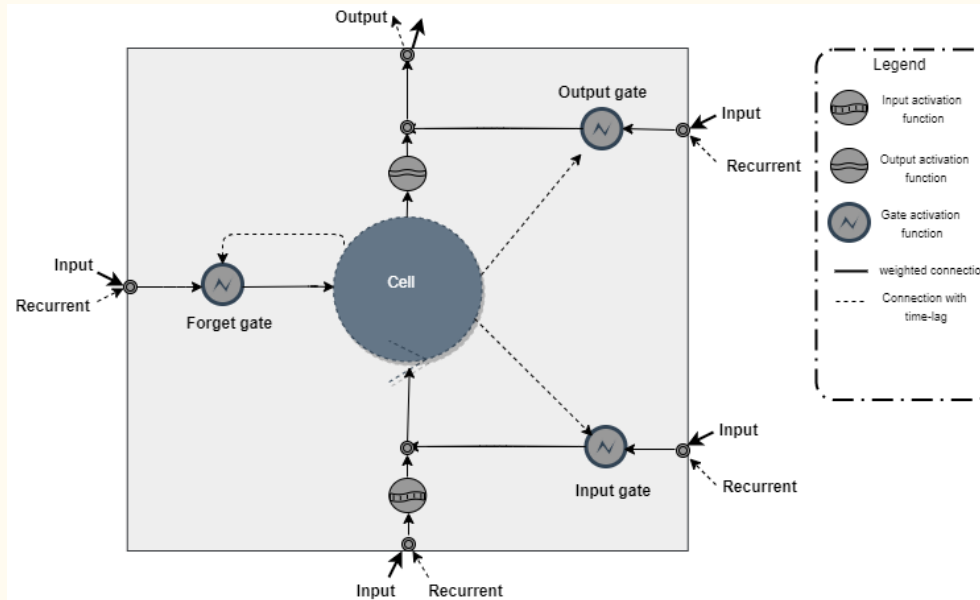Improvement of the recruitment process

KPI Evolution

CVs

SARIMA

$$SARIMA(p, d, q)(P, D, Q)^s$$

- p : the auto-regressive order
- d : the order of difference
- q : the order of the moving average
- P : the seasonal auto-regressive order
- D : the seasonal order of difference
- Q : the seasonal moving average
- s : the seasonal period

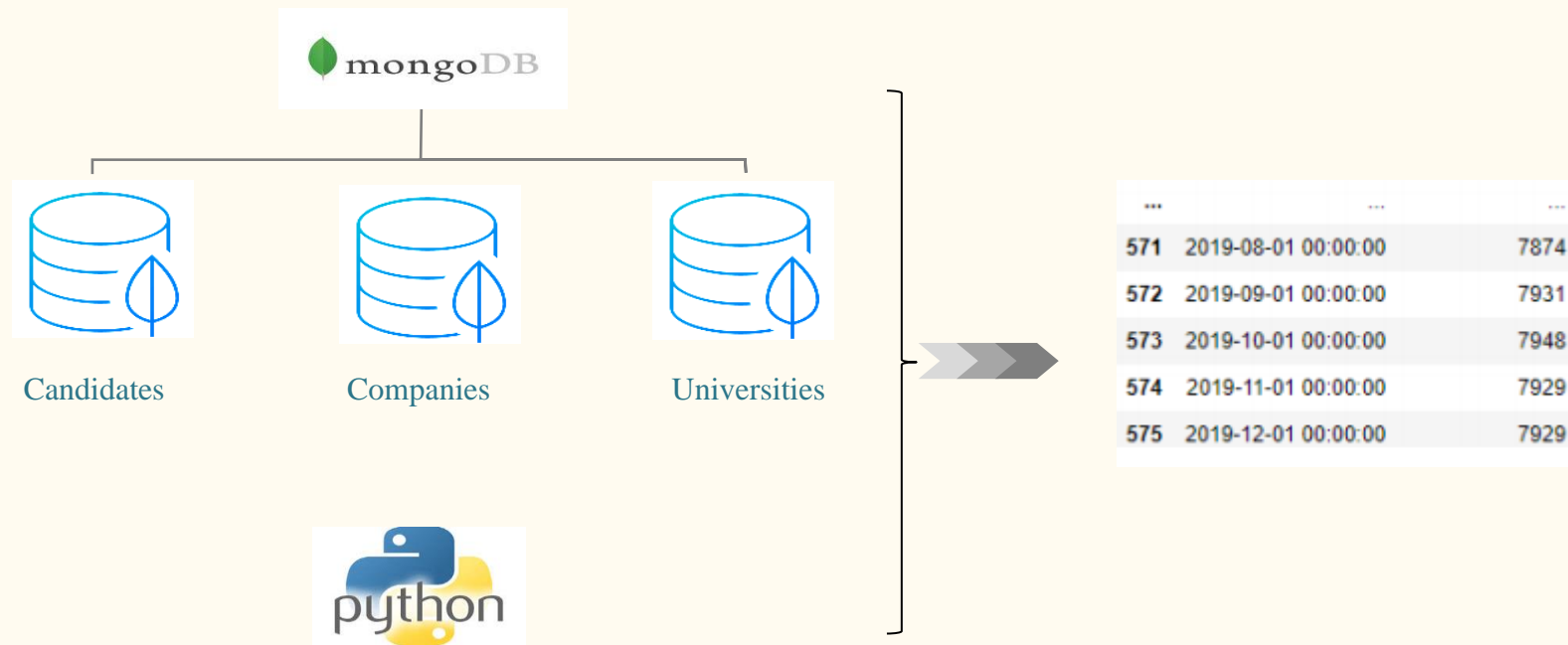$$\phi_P(L^s)\phi_P(L)(1-L)^d(1-L^s)^D X_t = \boldsymbol{\theta}_Q(L^s)\boldsymbol{\theta}_q(L)\boldsymbol{\varepsilon}_t$$

- $\phi_p \neq 0$, $\boldsymbol{\theta}_q \neq 0$, ϕ and $\boldsymbol{\theta}$ are two polynomials of respective degrees p and q
- $\boldsymbol{\varepsilon}_t$ is a white noise
- L is the lag operator defined by : $L(X_t) = X_{t-1}$

4

LSTM

Simple architecture of an LSTM block

Dataset

Candidates   Companies   Universities

...	...	...
571	2019-08-01 00:00:00	7874
572	2019-09-01 00:00:00	7931
573	2019-10-01 00:00:00	7948
574	2019-11-01 00:00:00	7929
575	2019-12-01 00:00:00	7929

Forecasting using SARIMA

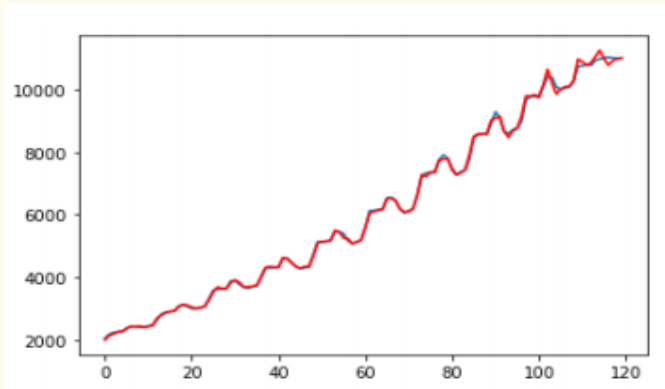Decomposition of time series:
example of number of candidates located in Tunisia
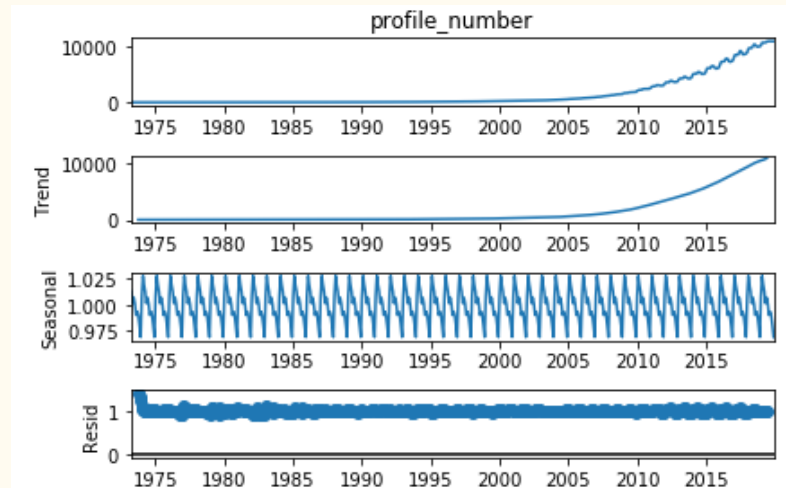
Forecasting using SARIMA

Number of candidates in Tunisia

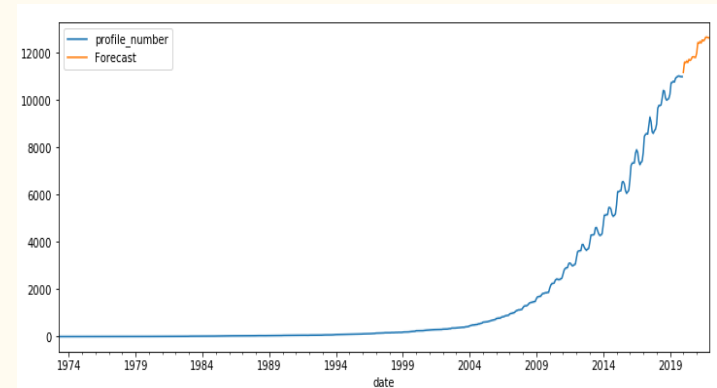SARIMA(1, 1, 2)(0, 2, 2, 12)

RMSE=73.04

Forecasting

Walk-forward validation

Predicted values vs actual values

8
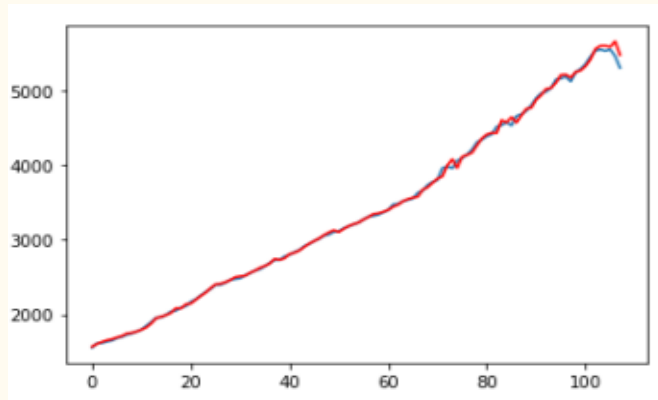
Forecasting using SARIMA

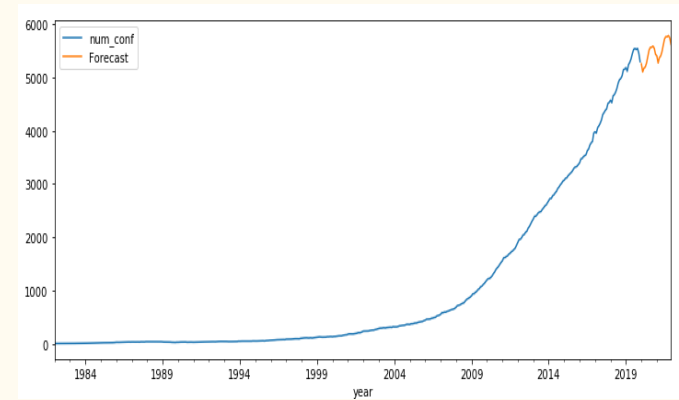## Number of confirmed candidates

SARIMA(2, 1, 2)(0, 2, 2, 12)

RMSE=41.27

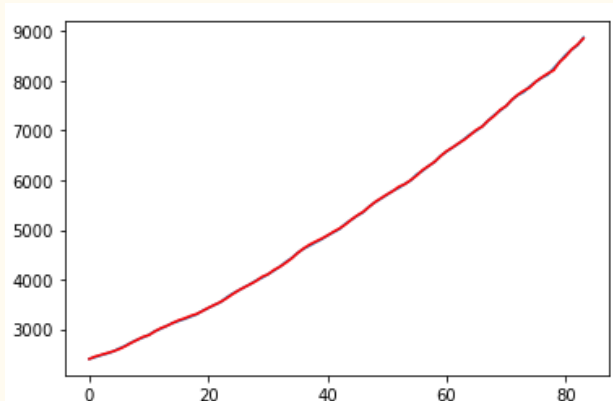Forecasting

Walk-forward validation

Predicted values vs actual values

9

Forecasting using SARIMA

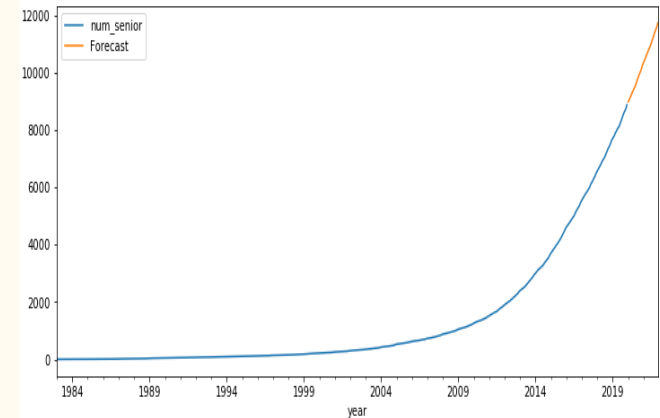## Number of senior candidates

SARIMA(1, 1, 1)(1, 1, 1, 12)

RMSE=12.86

Forecasting

Walk-forward validation

Predicted values vs actual values

10

Forecasting using LSTM

Number of candidates in Tunisia
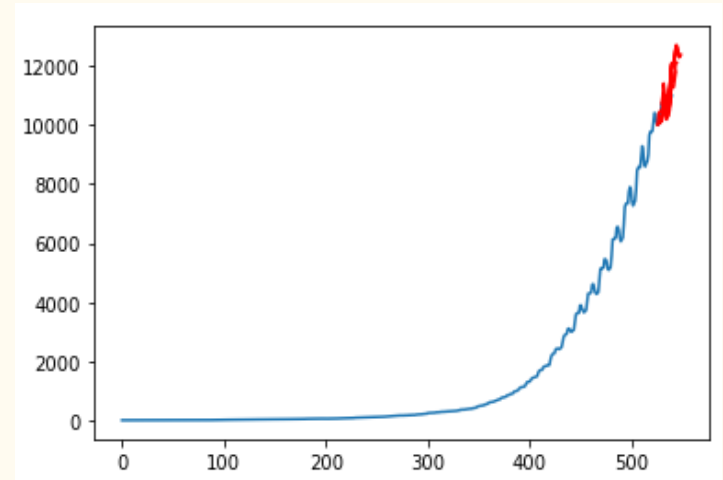
Hyperparameters

- n_lag = 12
- n_epochs = 100
- n_batch = 1
- n_neurons =50
- n_diff=1

Walk-forward validation

RMSE=268,5

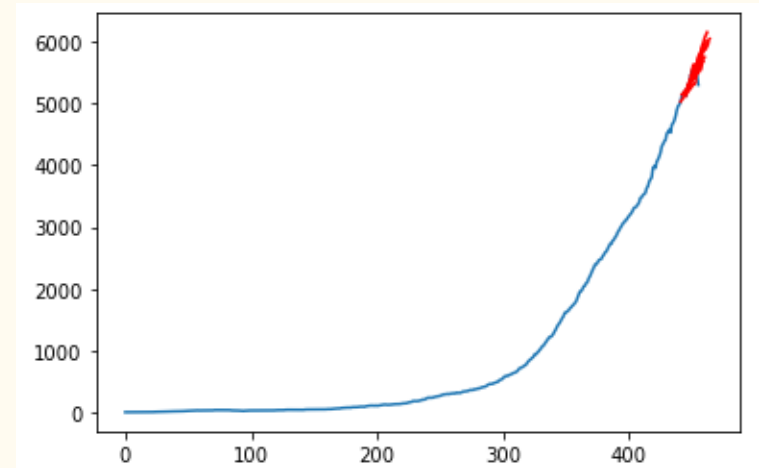# Forecasting using LSTM

Number of confirmed candidates

Hyperparameters

- n_lag = 12
- n_epochs = 1000
- n_batch = 1
- n_neurons =10
- n_diff=1

Walk-forward validation

RMSE=117,44
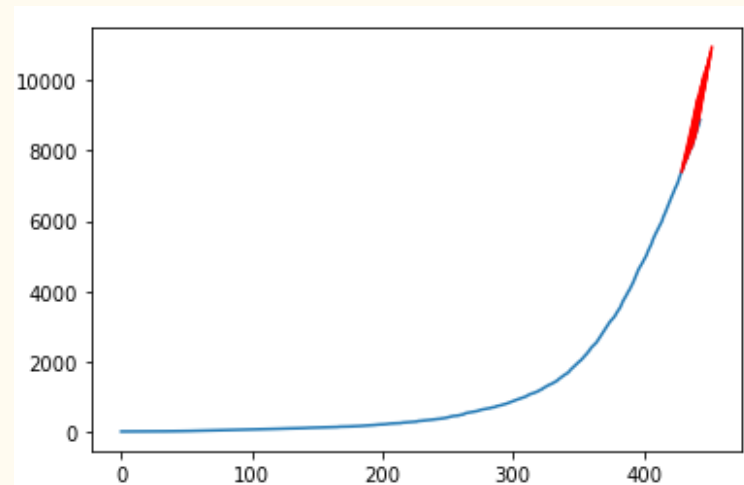
Forecasting using LSTM

Number of senior candidates

Hyperparamètres

- n_lag = 12
- n_epochs = 500
- n_batch = 1
- n_neurons =50
- n_diff=2

Walk-forward validation

RMSE=87,7

Conclusion

- SARIMA method was adapted to our case

- LSTM needed a lot of preparation to make data stationary

- Other LSTM network variants should be tested

- There are some other KPIs to be concerned as the number of candidates having a particular skill (reactjs, java …)

Thank you

Any questions ?

Feel free to contact me :

houichi.monia@gmail.com