IE 360 PROJECT REPORT

Contributors:

Ödül Tüfekçi 2018402180
Atıf Yönet 2018402174
Muhammet Talha Babur 2018402210

Introduction:

Our project is about providing hourly solar power prediction of KIVANC 2 GES (Güneş Enerjisi Santrali) for the next day by using the previous data. We tried two methods which are sarima and baseline method. We combined these methods manually and tried to find the best prediction.

Model Creation Phase:

1)Baseline Method:

Our first model is baseline method. Baseline method takes recent actual production as forecast. This model is quite inefficient in terms of not considering the previous data. Taking previous data into account gives us more accurate results. In addition to its disadvantages, it has also some advantages. In this kind of predictions which are affected by the season are predictions are similar to most recent data. This method is successful in predicting the night hours and hours which includes before sunrise. On the other hand, this method is quite unsuccessful in predicting hours between sunrise and sunset. Therefore, in our firstly forecast we used this method. After that, we want to try another method which is SARIMA.

Our Code for Baseline Method:

```
require(data.table)
require(forecast)
todays_date=Sys.Date()
```

```
forecast_date=todays_date+1
production_file=read.csv("productionnew.csv")
weather_file=read.csv("weathernew.csv")
production=fread("productionnew.csv")
weather=fread("weathernew.csv")
latest_available_prod_date=as.Date(max(production$date))
n_days=as.numeric(forecast_date-latest_available_prod_date)
forecasted_production=tail(production,n_days*24)
forecasted_production[,date:=date+n_days]
forecasted_production[,production:=NA]
production_with_forecast=rbind(production,forecasted_production)
forecast_table=data.table(date=forecast_date,hour=0:23,production=NA)
baseline_forecast=production_with_forecast[date==latest_available_prod_date]$production
forecast_table[,baseline:=baseline_forecast]
forecast_table

ggplot(production_with_forecast ,aes(x=date)) +
  geom_line(aes(y=production,color='real')) +
  geom_line(aes(y=production,color='predicted'))
```
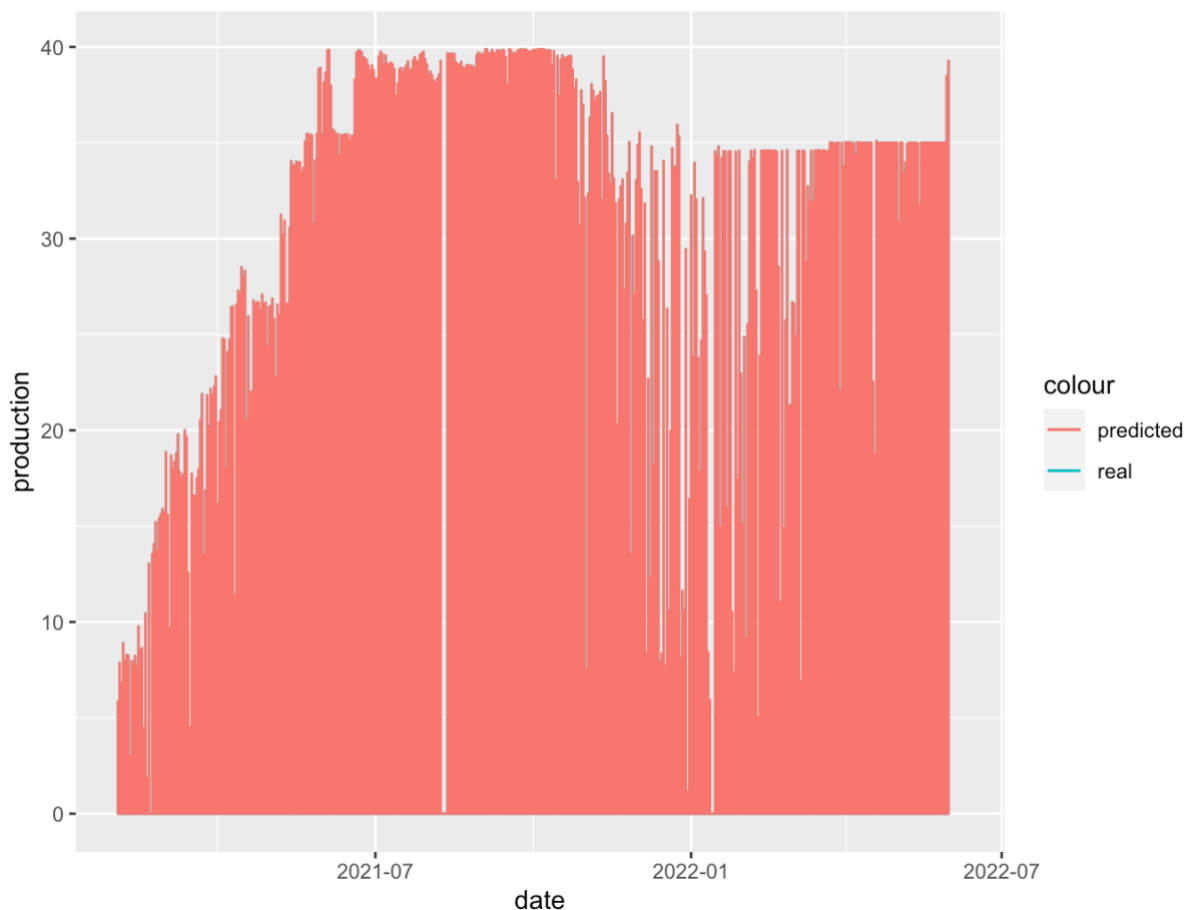

This code reads the excel file which is updated every day and we downloaded data everyday and change the file then using the most recent data as prediction.
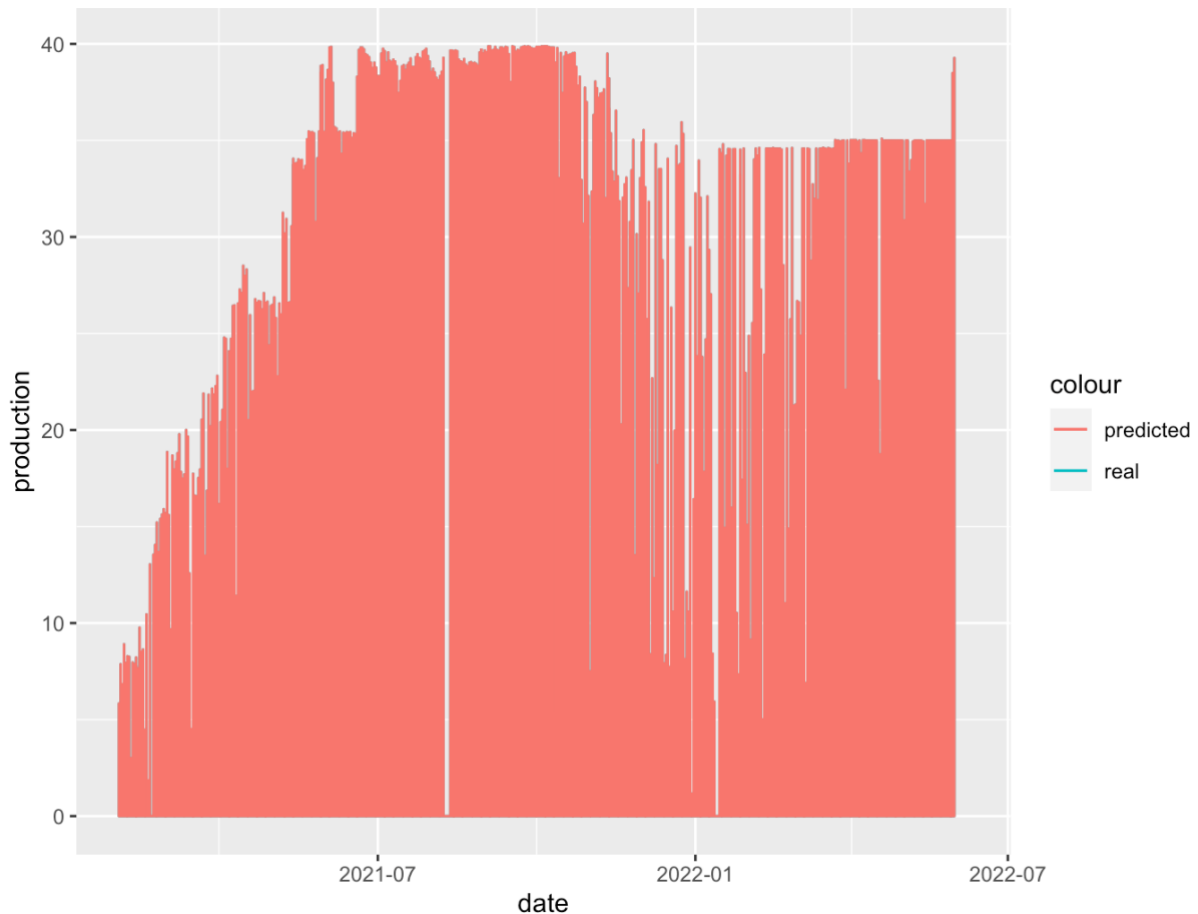
SARIMA Method:

Our second method is SARIMA method. SARIMA method has three new hyperparameters to specify the autoregression (AR), differencing (I) and moving average (MA) for the seasonal component of the series, as well as an additional parameter for the period of the seasonality. This method is more advantageous in terms of predicting hours between sunrise and sunset because this method takes previous data into account. However, this method is as not well as baseline method in terms of predicting the night hours and hours which includes before sunrise. Our code determines the best model using auto.arima function and find the best solution as ARIMA(3,0,1)(2,1,0).

Our Code for SARIMA Method:

```
require(data.table)
require(forecast)
todays_date=Sys.Date()
forecast_date=todays_date+1
production_file=read.csv("productionnew.csv")
weather_file=read.csv("weathernew.csv")
production=fread("productionnew.csv")
weather=fread("weathernew.csv")
latest_available_prod_date=as.Date(max(production$date))
n_days=as.numeric(forecast_date-latest_available_prod_date)
forecasted_production=tail(production,n_days*24)
forecasted_production[,date:=date+n_days]
forecasted_production[,production:=NA]
production_with_forecast=rbind(production,forecasted_production)
forecast_table=data.table(date=forecast_date,hour=0:23,production=NA)
production_with_forecast=production_with_forecast[order(date,hour)]
production_series=ts(production_with_forecast[!is.na(production)]$production,frequency=24)
sarima_model=auto.arima(production_series,seasonal=T,stepwise=T,approximation=T,trace=T)
forecast_ahead=nrow(forecast_table)
sarima_forecast=forecast(sarima_model,h=forecast_ahead)
forecast_table[,sarima:=tail(sarima_forecast$mean,24)]
forecast_table

ggplot(production_with_forecast ,aes(x=date)) +
  geom_line(aes(y=production,color='real')) +
  geom_line(aes(y=production,color='predicted'))
```

This code reads the excel file which is updated every day and we downloaded data every day and change the file then SARIMA code takes frequency as 24 and predicted the results.

Conclusion:

In this project, we combine two methods which are SARIMA and baseline method. In competition phase first, we tried these methods separately. Then, we combine these methods manually in order to gain both method's benefits and prevent their unique disadvantages. We use baseline method for predicting hours between sunset and sunrise and use SARIMA method for predicting hours between sunrise and sunset.