



GREEN ANALYTICS

Bence Hervay, Jakob Alwall, Junhua Chen,
Kornél Szabó

MOTIVATION: GREEN FINANCE

- Carbon Allowances: tradable illiquid asset allowing an organisation to emit carbon
- Aim to Increase liquidity in carbon allowances
 - Create value and reduce risk for ordinary people
 - Create upward pressure on price of carbon allowances
 - Incentivize de-carbonisation



```
6, 82.08, 80.58, 82.67, 82.09, 85.59, 84.47] 73
```

```
22]: headlines_df = pd.read_csv("climate-news-db-dataset.csv")
```

```
23]: headlines_df
```

```
23]:
```

	newspaper_id	newspaper	newspaper_url	headline	article_url	article_id	date_p
0	atlantic	The Atlantic	theatlantic.com	Why Arugula Is the Best Leafy Green	https://www.theatlantic.com/health/archive/201...	585571	25T
1	atlantic	The Atlantic	theatlantic.com	Scott Pruitt Rejects the Scientific Fact of Cl...	https://www.theatlantic.com/science/archive/20...	519054	09T
2	atlantic	The	theatlantic.com	None	https://www.theatlantic.com/national/archive/2...	258549	

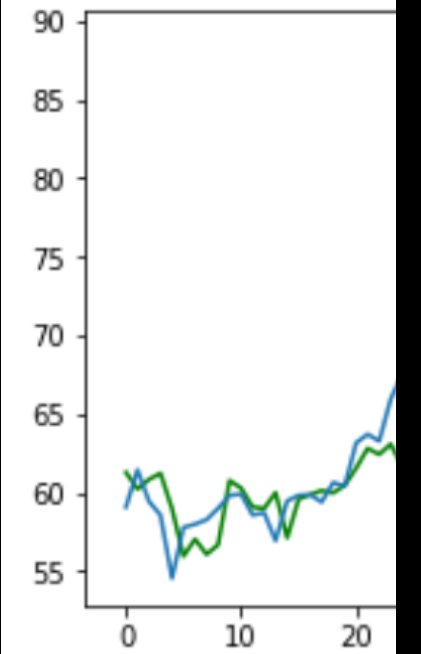
```
1456 280.702033 279.563496 282.075691 81.211382 0.000000 0.0 12.0 0 684773.0
1457 281.067398 279.897967 282.467073 82.593496 0.002439 0.0 12.0 1 687306.0
1458 281.231583 280.291667 282.350000 78.708333 0.007500 0.0 12.0 2 639282.0
1459 280.159274 278.972581 281.601613 74.451613 0.000000 0.0 12.0 3 606334.0
1460 280.438264 279.290496 281.745041 70.595041 0.000000 0.0 12.0 4 621534.0
```

```
1443 rows x 9 columns
```

```
from sklearn.model_selection import train_test_split
# Separate target from predictors
y = final.answers
X = final.drop(columns=['answers'])
# Divide data into training and validation subsets
# 80-20 split with random shuffling
X_train_full, X_valid_full, y_train, y_valid = train_test_split(X, y, train_size=0.8, test_size=0.2,
random_state=0)
```

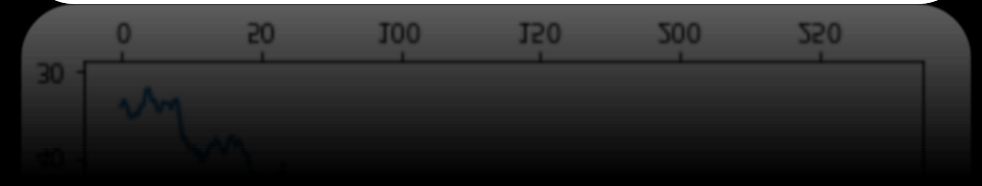
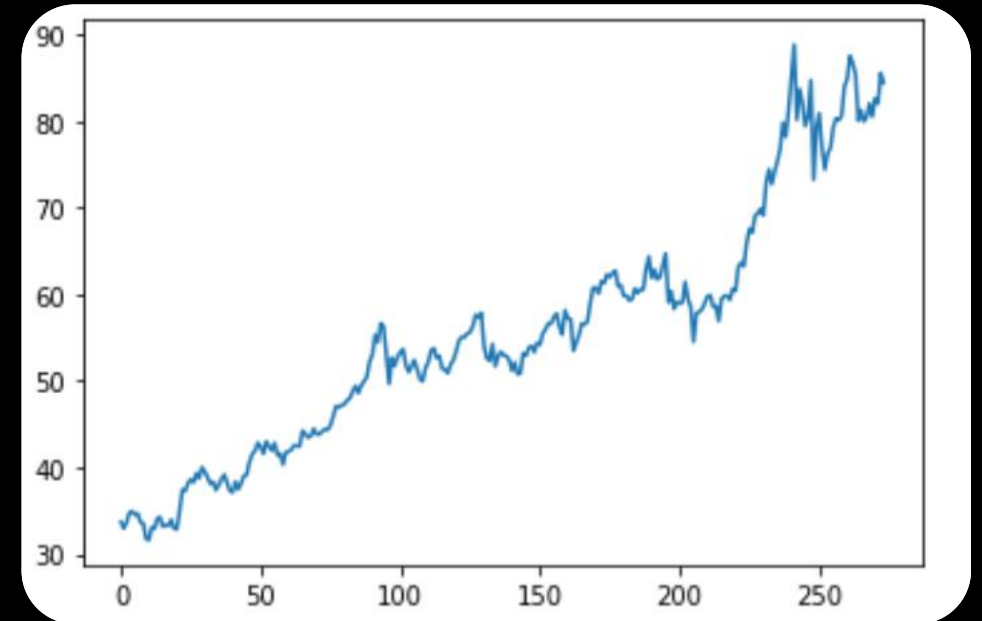
```
gbModel = XGBRegressor
gbModel.fit(X_train
plt.plot(gbModel.pr
plt.plot(y_test)
```

```
[<matplotlib.lines.
```



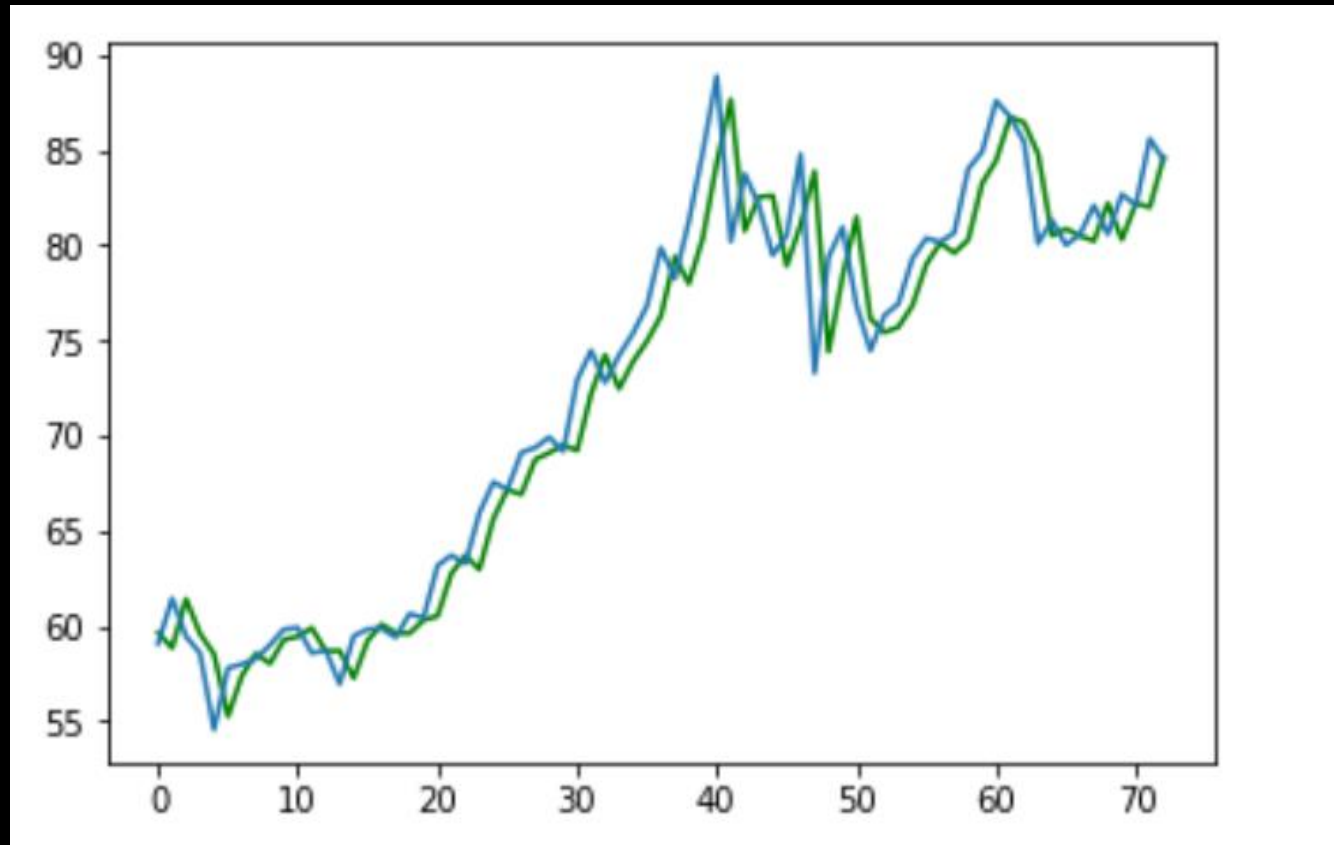
TIME SERIES FORECASTING OF EMISSION ALLOWANCE PRICES

- Price forecasts are useful
- Real time updates by scraping



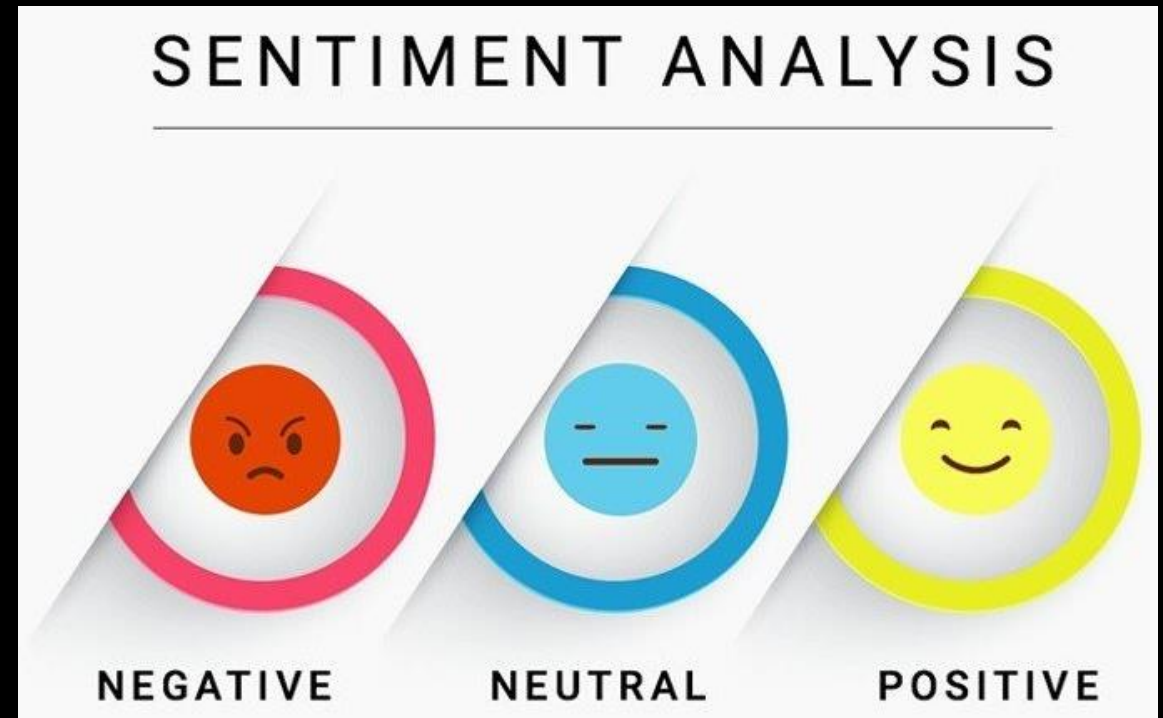
HISTORICAL REGRESSION

Validation Set performance (green = forecast, blue = ground truth)



SENTIMENT CLASSIFICATION

- Consumer sentiment & News events is key to forecasting price movements.
 - A greater sense of urgency on climate change makes carbon emissions more costly, evidenced by increase in emissions prices
- Climate fuelled disasters bring grave costs across globe 2021 (greater urgency)
- For the first time, electric and hybrid vehicles outsold diesel vehicles in Europe positive (less urgency, signals decrease in demand for CO2 emission)



TRANSFER LEARNED GPT-3 MODEL

- 88% accuracy on self-labelled dataset, classifying news articles as either positive or negative
- Scrape news from Reuters and classify sentiment, then aggregate

```
advisers slam eu plan label gas nuclear investments green draft -1  
lead eu lawmaker wants airlines pay their co2 emissions sooner 1  
exxon pledges net zero carbon emissions operations by 2050 1  
how white house hopes save bidens spending bill 1
```

Sample predictions on recently scraped Reuters Climate
News titles

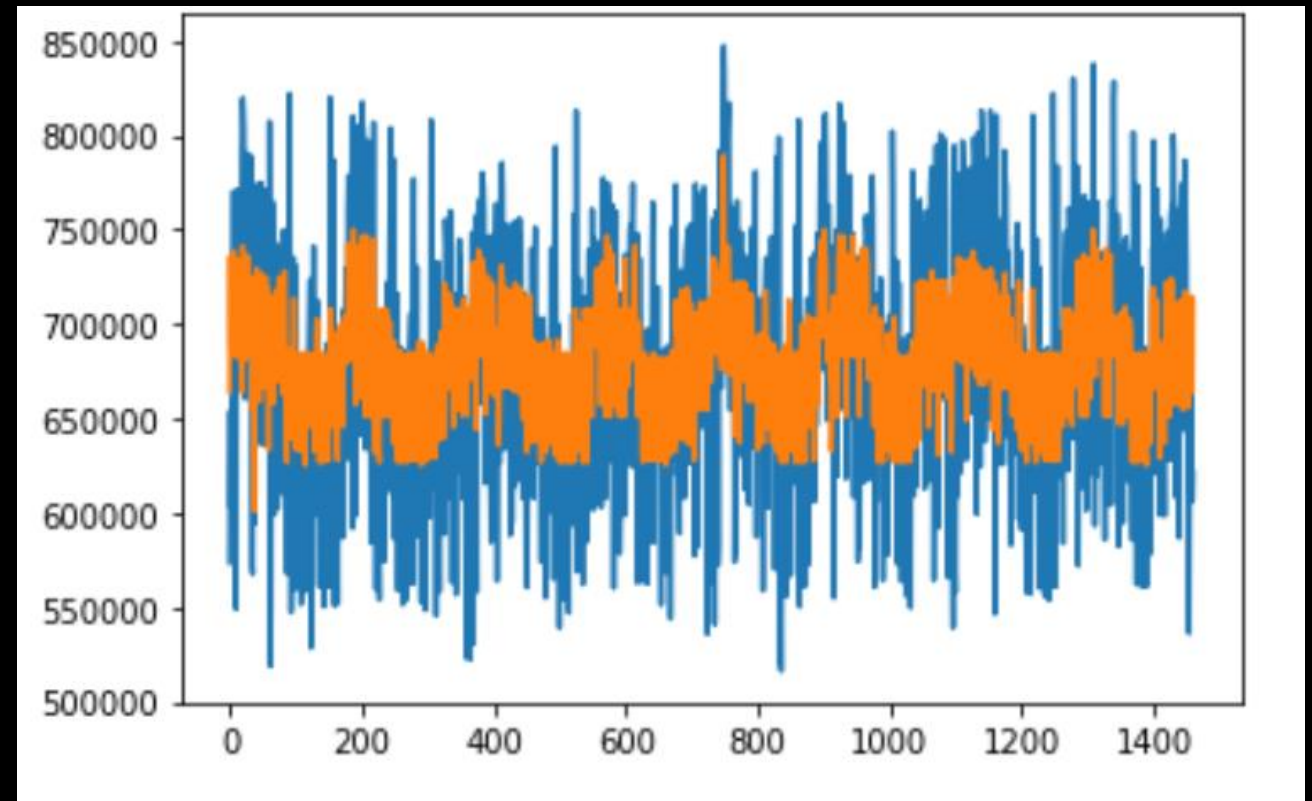
ELECTRICITY DEMAND PREDICTOR

- Demand for electricity can be predicted to good accuracy by considering factors like time of year and weather
- Demand for electricity correlates strongly with demand for emissions, thus is a factor in pricing



XGBOOST MODEL TRAINED ON SPANISH ENERGY DEMANDS

- Feature engineering, Data cleaning
- Nicely captures major elements in predictions



Prediction

Ground truth