Predicting the Price of Oil

Team #10

		
Bogdana Kolic	Clemence Mottez	Matheo Le Masson
• Model selection • Data preprocessing	• Feature selection • Data analysis	• PCA • Data collection

Highlights:

- The bests model to predict oil price is XGB regressor
- The 5 most important features to predict the price of oil are:
  - 1) Year
  - 2) World imports
  - 3) Inflation
  - 4) Price of gold
  - 5) War (geopolitical conflicts involving OPEC countries)
- With the selected features, we obtain a lower MSE
- PCA components provide a slightly better score than our selection but we lose in interpretability if we use them

Brief Report:

For this project we have selected **23 parameters** ranging from economic indicators to geopolitical factors, and other relevant variables. We then ran and compared multiple models on all this data. **XGB regressor** is the model who stood out as having the lowest **MSE**. We used **PCA** analysis to reduce the dimension of our data, and found that we needed 5 principal components to explain 90% of the variance, and that added a sixth did not change much. We then did a **selection of 5 features** using various methods: filter models, wrapper models and embedded models. Thus, we decided to **compare** our selection of 5 features to PCA to see which one worked best. We find that our selection improves the score obtained by several models, and while it is a bit less accurate than PCA components, it provides **interpretability**.

Description of Appendices (Optional):

See github: https://github.com/Oil-Price-Prediction/Oil-Price-Prediction