Using ChatGPT to predict the Equity Markets

E. Benhamou, B. Guez, JJ. Ohana, D. Saltiel, B. Lefort

For Alpha

Dauphine | PSL UNIVERSITÉ PARIS

LAMSADE UMR 7243

PSL Week
AI on
Economics & Finance

Thursday
30 Nov 2023

E. Benhamou

Your instructor

- Mix of experience between private and public research and startups

Goldman Sachs

NATIXIS

REFINITIV

For Alpha

- Teaching

Dauphine | PSL UNIVERSITÉ PARIS

LAMSADE UMR 7243

- Combined expertise: 3 PhDs in Economics, Maths and Computer sciences
- Top contributor to

SSRN

1. NLP

What is NLP?

- Natural Language Processing (NLP) is a field of computer science, artificial intelligence, and linguistics concerned with the interactions between computers and human (natural) languages.
  - **Human-Computer Interaction:** NLP enables computers to understand, interpret, and respond to human language.
  - **AI and Linguistics Fusion:** It combines artificial intelligence with linguistic principles.
  - **Data Processing:** NLP processes large volumes of natural language data.
  - **Language Analysis:** It involves analyzing the structure and meaning of language.
  - **Diverse Applications:** NLP is used in various technologies like voice assistants, translation services, and sentiment analysis.

Why NLP algorithms are difficult?

- **Ambiguity and Nuances**: NLP algorithms struggle with the nuances and ambiguity inherent in human language.
- **Language Diversity**: The wide variety of languages and dialects adds complexity.
- **Context Understanding**: Algorithms must interpret context, a challenging task.
- **Figurative Language**: Identifying sarcasm and figurative speech is difficult.
- **Evolving Language**: Continuous language evolution demands frequent updates.

Ambiguity and specificity of financial language.

- **Specialized Jargon:** Financial language includes specific jargon and technical terms.
- **Context-Dependent Meanings:** Terms can have multiple meanings based on context.
- **Nuanced Expressions:** Subtleties in language can alter the meaning significantly.

- *Example:* Bullish -> does not specify which market or asset and the degree of optimism

Speed of financial markets.

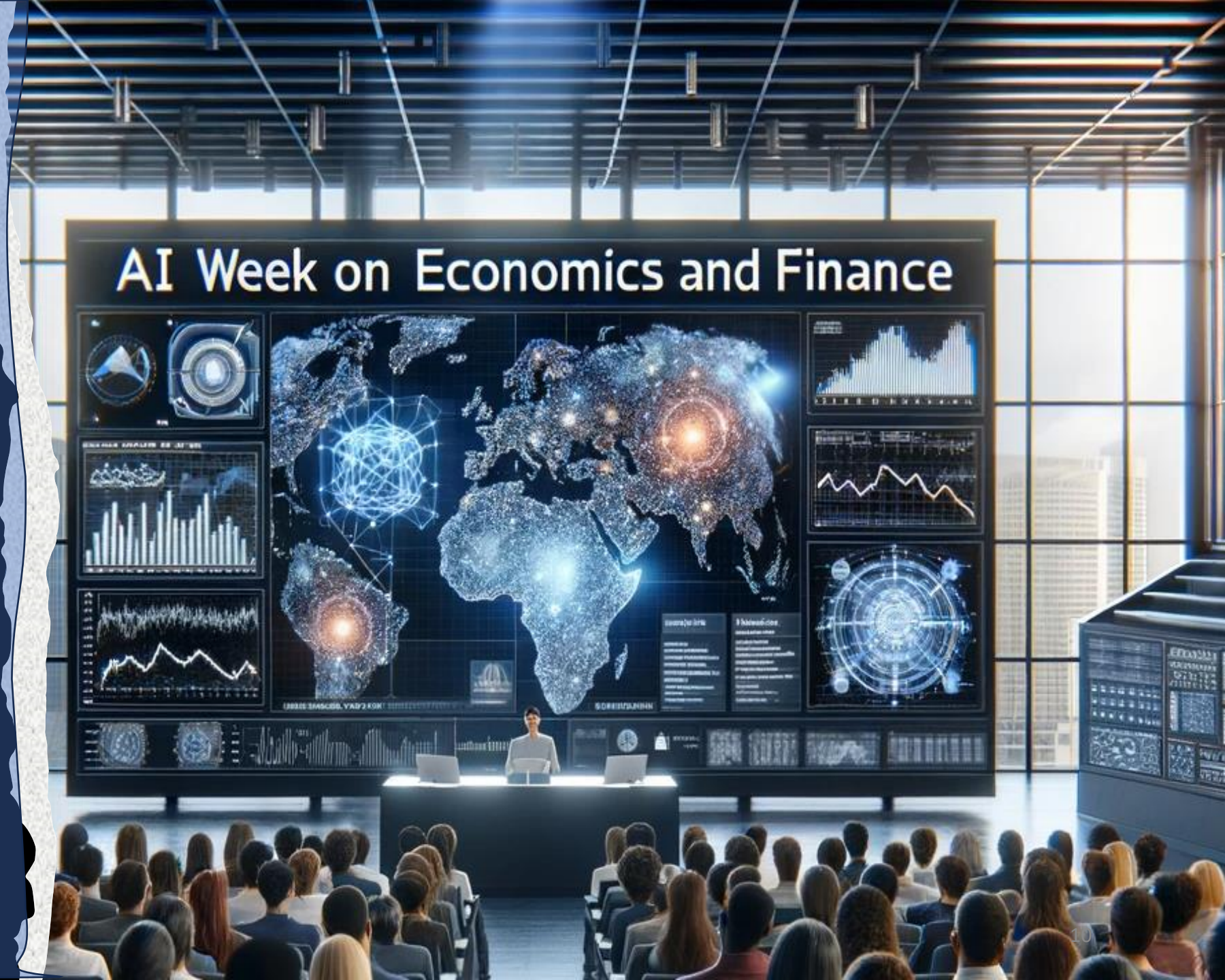
- **Rapid Transactions:** Financial markets execute transactions at high speeds, often in milliseconds.
- **Real-Time Data Flow:** Markets continuously generate vast amounts of real-time data.
- **Quick Reaction to News:** Financial markets rapidly react to global news and economic events.

But at the end

- Finally, we can conclude that NLP in finance is complex because of language complexity:

- **Language Complexity:** The financial language is filled with jargon, ambiguous expressions, and metaphors, making it challenging for NLP algorithms to interpret.
- **Subtle Nuances:** Financial reports and market analyses often contain subtle nuances that are difficult for NLP algorithms to correctly understand and process.

2. Sentiment Indicators



Concepts

- **Emotional Analysis**: A sentiment indicator measures the overall emotional tone in written or spoken language, often categorizing it as positive, negative, or neutral.

- **Market Insight Tool**: In finance, it's used to gauge investor sentiment and market mood, providing insights into market trends and investor behavior.

- **Data Source Variety:** These indicators analyze various data sources, like social media, news articles, and financial reports, to assess public sentiment towards specific assets or the market as a whole.

In practice …

- **Polarity Score:**
  - Measures sentiment from -1 (negative) to 1 (positive).
  - Can be very simple to complex

- **TF-IDF (Term Frequency-Inverse Document Frequency):**
  - While not a direct sentiment analysis method, TF-IDF is often used in preprocessing to identify significant terms in a document

- **Lexicons:**
  - Lexicons like SentiWordNet or AFINN assign predefined sentiment scores to words. The sentiment of a text is determined by aggregating these scores.

Polarity score

- **Polarity Score:**
  - Measures sentiment from -1 (negative) to 1 (positive).

$$\text{Polarity Score} = \frac{\text{Number of Positive Words} - \text{Number of Negative Words}}{\text{Total Number of Words}}$$

- **What are the challenges?**

Challenges in polarity score

- **Contextual Meaning:** Words may have different meanings based on context, making it hard for algorithms to accurately assess sentiment.
- **Sarcasm and Irony:** Detecting sarcasm or irony is complex, as they often imply the opposite of the literal word meaning.
- **Neutral Phrasing:** Neutral or factual statements can be misinterpreted as positive or negative, skewing the polarity score.

TF-IDF (Term Frequency-Inverse Document Frequency)

- This is used to **weigh the importance of words** in a document set. It increases **proportionally to the number of times a word** appears in the document but is offset by the frequency of the word in the corpus. The formula is:

$$\text{TF-IDF} = \text{TF}(t, d) \times \text{IDF}(t, D)$$

where

- **TF(t,d)** is the term frequency, and
- **IDF(t,D)** is the inverse document frequency.

VADER (Valence Aware Dictionary and sEntiment Reasoner)

- VADER is a lexicon and rule-based sentiment analysis tool specifically attuned to sentiments expressed in social media.
- It provides a compound score that aggregates the sum of all lexicon ratings, normalized between -1 (most negative) and +1 (most positive).

VADER in practice

```
!pip install vaderSentiment

from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
# initialize the sentiment analyzer
sa = SentimentIntensityAnalyzer()
```

VADER in practice

```
# The polarity score() method returns a float for the sentiment strength based on the input text


statements = [
    "What a wonderful day",
    "Thank you for clarifying that my eye cancer will not yield me deaf. I consider myself extremely fortunate that an intellectual giant such as yourself would choose to operate on me",
    "We shared our sadness at the waste of two barely emerging lives with the remainder of the celebratory bourbon.",
    "She didn't know whether to be angry or frightened because she had never been propositioned before."
]


for s in statements:
    scr = sa.polarity_scores(s)["compound"]
    print(f'Sentiment value of:-\n"{s}" is :- {scr}')
```

3. News horizons

News impact

- **Minutes (Intraday Trading):** Immediate news is crucial for intraday traders who use minute-by-minute updates to capitalize on quick market movements for short-term gains.

- **Days to Weeks (Short-Term Strategies):** News within a daily to weekly timeframe guides short-term traders, helping them leverage slightly longer market trends while maintaining flexibility in their strategies.

- **Longer Horizons (Long-Term Strategies):** Long-term investors focus on news with lasting impacts, unfolding over months or years, to align their portfolios with broader economic trends and long-term goals.

For which time horizons? 1/2

- **Immediate News (Seconds to Minutes):**
  - Ideal for high-frequency and algorithmic traders.
  - Relies on real-time data for quick decision-making and exploiting short-lived market opportunities.
- **Short-Term News (Hours to Days):**
  - Suitable for day traders and swing traders.
  - Focuses on daily market trends, earnings reports, and economic announcements.

For which time horizons? 2/2

- **Medium-Term News (Weeks to Months):**
  - Beneficial for position traders.
  - Involves analyzing weekly or monthly market trends, policy changes, and broader economic indicators.
- **Long-Term News (Months to Years):**
  - Important for buy-and-hold investors and long-term portfolio managers.
  - Focuses on fundamental analysis of industries, long-term economic trends, and significant global events.

4. Using news

How to measure an investment strategy

- Risk-Adjusted Returns: The **Sharpe Ratio** measures the performance of an investment compared to a risk-free asset, after adjusting for its risk and is given by:

$$\text{Sharpe Ratio} = (R_p - R_f)/ \sigma$$

Where:

- R_p is the return of the portfolio, R_f is the risk-free rate, and σ the standard deviation of the portfolio's excess return.
- Interpretation: A higher Sharpe Ratio indicates better risk-adjusted returns. It's used to compare the performance of different investments or portfolios.

Achievable Sharpe Ratio 1/2

- **Typical Range**: In most investment scenarios, the Sharpe Ratio typically ranges between 0.5 to 1.5. This range is considered indicative of good to excellent performance.

- **High-Frequency Trading Exception:** High-frequency trading strategies can achieve higher Sharpe Ratios, often exceeding the typical range, due to their ability to capitalize on very short-term market inefficiencies.

Achievable Sharpe Ratio 2/2

- **Risk and Reward Balance:** The Sharpe Ratio reflects the balance between risk and reward. A ratio between 0.5 to 1.5 suggests a reasonable return for the level of risk taken.

- **Dependent on Market Conditions:** Achievable Sharpe Ratios can vary significantly depending on market conditions, investment strategy, and the risk-free rate used in the calculation.

- **Benchmarking Tool:** Investors use the Sharpe Ratio to benchmark and compare the risk-adjusted performance of different portfolios or assets.

5. ChatGPT

Introduction to ChatGPT

- **AI Language Model:** Developed by OpenAI, based on the GPT (Generative Pre-trained Transformer) architecture.
- **Deep Learning:** Utilizes advanced machine learning techniques for natural language understanding and generation.
- **Interactive Communication:** Designed to converse with users, providing responses that mimic human-like text.

Capabilities of ChatGPT

- **Diverse Conversations:** Engages in various topics, from general knowledge to specific queries.
- **Language Skills:** Capable of understanding and generating text in multiple languages.
- **Learning Ability:** Continuously learns from interactions to improve response quality.

Applications of ChatGPT

- **Wide Applications:** Used in customer service, content creation, education, and more.
- **Customization:** Adaptable to different industries and user needs.

- Because it is used without learning, it is called **zero shot learning**

Using ChatGPT in practice 1/5

```
!pip install openai

import os
import openai
# set openai api_key
api_key = 'YOUR API KEY'
os.environ['OPENAI_API_KEY'] = api_key

# initialize
from openai import OpenAI
client = OpenAI(api_key=os.environ['OPENAI_API_KEY'])
```

Using ChatGPT in practice 2/5

```
# create an ansynchronous call
from openai import AsyncOpenAI
client = AsyncOpenAI()
completion = await
client.chat.completions.create(model="gpt-3.5-
turbo", messages=[{"role": "user", "content":
"Hello world"}])
```

Using ChatGPT in practice 3/5

```
print(completion)
```

```
> ChatCompletion(id='chatcmpl-8QNMNOONdot2cLyGOuVg7EyO9UvlS',
choices=[Choice(finish_reason='stop', index=0,
message=ChatCompletionMessage(content='Hello! How can I assist you today?',
role='assistant', function_call=None, tool_calls=None))], created=1701297015,
model='gpt-3.5-turbo-0613', object='chat.completion', system_fingerprint=None,
usage=CompletionUsage(completion_tokens=9, prompt_tokens=9, total_tokens=18))
```

```
print(completion.choices[0].message.content)
```

```
Hello! How can I assist you today?
```

Using ChatGPT in practice 4/5

```
json_result = openai.Image.create(
      prompt="Create an image of the PSL week about
using AI for Economics and Finance",
       n=2, size="1024x1024")
```

> {'created': 1684475787454, 'data': [{'url':
'https://files.oaiusercontent.com/file-zaZ7KrktVvpYdRuQYwJoPrB0?se=2023-11-29T23%3A51%3A08Z&sp=r&sv=2021-08-06&sr=b&rscc=max-age%3D31536000%2C%20immutable&rscd=attachment%3B%20filename%3D13a9cbd9-3cfb-4efa-8ea6-c1bd9deb8679.webp&sig=JHGxryx6kwx0CDzYDNnvVKgbXIjJUbEsFYgvnKhMJTA%3D'},
{'url': 'https://files.oaiusercontent.com/file-GQ0oAWz0VQn9Xtb8ADquF63a?se=2023-11-29T23%3A51%3A08Z&sp=r&sv=2021-08-06&sr=b&rscc=max-age%3D31536000%2C%20immutable&rscd=attachment%3B%20filename%3Dab1348ce-80d2-483a-b5d2-5faa7598238c.webp&sig=DXzHp3Vj5bMA31%2BbYdVdn66fG0/VBI3bF7Ncg9aKzvQ%3D'}]}

Using ChatGPT in practice 5/5

```
from IPython.display import Image
# Replace the URL below with the URL of your image
for img in  json_result['data']:
  display(Image(url=img['url']))
```

Resulting images

Google Collab notebook

- https://colab.research.google.com/drive/1yRj2cH5MjW6hnwD5_PW:cQ?usp=sharing

6. Comparing with other NLP models

What is FinBERT?

- **Specialized AI Model:** FinBERT is a variant of the BERT (Bidirectional Encoder Representations from Transformers) model, fine-tuned for financial contexts.
- **Finance-Specific Training:** Trained on a large corpus of financial texts for nuanced understanding of financial language.
- **Sentiment Analysis:** Excellently equipped for analyzing sentiments in financial news and reports.

How does this differs from ChatGPT?

- **Versatile Language Model:** Developed by OpenAI, based on the GPT architecture, suitable for a broad range of topics.
- **Conversational Expertise:** Specializes in generating human-like text and engaging in diverse conversations.
- **Adaptive Learning:** Continuously updates its knowledge and response quality through interactions.

Comparing FinBERT and ChatGPT

- **Specialization vs. Versatility:** FinBERT is specialized for financial text analysis, while ChatGPT is versatile across numerous subjects.
- **Sentiment Analysis vs. Conversational Ability:** FinBERT excels in sentiment analysis in finance, whereas ChatGPT is adept at conversational interactions in various domains.

=> So to use ChatGPT, rely on its capacity to make summary and do a **one shot learning.**

7. Correlation Measures

Pearson Correlation Coefficient

- **Definition**: Measures the linear relationship between two variables.
- **Scale**: Ranges from -1 to +1, where +1 indicates a perfect positive linear relationship, -1 indicates a perfect negative linear relationship, and 0 implies no linear correlation.
- **Calculation**: Based on the mean and standard deviation of the variables. Formula for two times series x and y with obvious notations

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Spearman Correlation Coefficient

- **Definition:** Assesses how well the relationship between two variables can be described using a monotonic function.
- **Non-Parametric:** Does not assume a normal distribution of the data.
- **Calculation:** Based on the ranked values of the data rather than the raw data. Used when the relationship is not linear or the data is ordinal.

Spearman correlation calculation

For a sample of size *n*, the *n* raw scores X_i, Y_i are converted to ranks $\mathrm{R}(X_i), \mathrm{R}(Y_i)$, and r_s is computed as

$$r_s = \rho_{\mathrm{R}(X),\mathrm{R}(Y)} = \frac{\mathrm{cov}(\mathrm{R}(X), \mathrm{R}(Y))}{\sigma_{\mathrm{R}(X)} \sigma_{\mathrm{R}(Y)}},$$

where

ρ denotes the usual Pearson correlation coefficient, but applied to the rank variables,

$\mathrm{cov}(\mathrm{R}(X), \mathrm{R}(Y))$ is the covariance of the rank variables,

$\sigma_{\mathrm{R}(X)}$ and $\sigma_{\mathrm{R}(Y)}$ are the standard deviations of the rank variables.

Spearman correlation Property

Only if all *n* ranks are *distinct integers*, it can be computed using the popular formula

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)},$$

where

$d_i = \mathrm{R}(X_i) - \mathrm{R}(Y_i)$ is the difference between the two ranks of each observation,

n is the number of observations.

Pearson vs. Spearman Correlation

- **Nature of Relationship:** Pearson measures linear relationships, while Spearman measures monotonic relationships (either increasing or decreasing, but not necessarily linear).
- **Data Sensitivity:** Pearson is sensitive to outliers, whereas Spearman is more robust to outliers due to its use of ranks.
- **Data Type Suitability:** Pearson is ideal for continuous and normally distributed data; Spearman is better for ordinal data or non-normal distributions.

9. Time series

Understanding Time Series Cross-Validation

- **Special Nature of Time Series Data:** Time-dependent, sequential nature makes traditional cross-validation techniques inappropriate.
- **Objective:** To validate model performance while maintaining the temporal order of observations.

Key Considerations in Time Series

- **Temporal Order:** Ensuring that the model is not trained on future data.
- **Trend and Seasonality:** Accounting for potential trends and seasonal effects in the data.
- **Data Leakage:** Preventing information from future data points leaking into the model training process.

Time Series Cross-Validation Techniques

- **Walk-Forward Validation:** Sequentially stepping forward in time, expanding the training dataset and testing on subsequent data.
- **Rolling-Window Analysis:** Using a moving 'window' of data for training, testing on the following window, then rolling the window forward.
- **Time Series Split:** Dividing the time series into distinct training and testing periods to evaluate model stability over time.

Implementing Cross-Validation

- **Data Partitioning:** Carefully partition data in time order, ensuring no overlap between train and test sets.
- **Evaluation Metrics:** Choosing appropriate metrics (e.g., MAE, RMSE) relevant to time series forecasting.
- **Iterative Testing:** Repeatedly testing the model on different time periods to assess performance consistency.

Best Practices

- **Model Complexity:** Be wary of overfitting, especially in time series with many parameters.
- **External Factors:** Consider external variables that may affect the time series (e.g., economic indicators).
- **Reproducibility:** Ensure that the method and results are reproducible with different sets of data.

10. Overfitting

Introduction to Overfitting

- **Definition:** When a model learns the training data too well, including noise and outliers, and performs poorly on new data.
- **Symptoms:** Excellent performance on training data but poor generalization to new, unseen data.
- **Risk:** Overfitting leads to models that are not useful for making predictions on real-world data.

Data Management Strategies

- **Data Quality:** Ensure clean, relevant, and diverse data for training.
- **Data Augmentation:** Increase the size and variety of the training set to improve model robustness.
- **Feature Selection:** Remove irrelevant or redundant features to simplify the model.

Model Complexity and Regularization

- **Simplifying the Model:** Choose the right model complexity to fit the data adequately.
- **Regularization Techniques:** Implement L1 or L2 regularization to penalize overly complex models.
- **Cross-Validation:** Use cross-validation to test the model's ability to generalize to new data.

Training Techniques

- **Early Stopping:** Stop training when performance on a validation set starts to degrade.
- **Learning Rate Adjustment:** Optimize the learning rate to prevent rapid convergence to local minima.
- **Batch Size Optimization:** Experiment with different batch sizes for stable and effective learning.

Best Practices and Monitoring

- **Ensemble Methods:** Use techniques like bagging or boosting to improve model generalization.
- **Performance Metrics:** Monitor both training and validation metrics to detect overfitting.
- **Continuous Evaluation:** Regularly evaluate the model against new data to ensure it remains relevant and accurate.

Best Practices and Monitoring

- **Ensemble Methods:** Use techniques like bagging or boosting to improve model generalization.
- **Performance Metrics:** Monitor both training and validation metrics to detect overfitting.
- **Continuous Evaluation:** Regularly evaluate the model against new data to ensure it remains relevant and accurate.

Your time to play

- https://www.kaggle.com/competitions/2023-psl-week-using-bloomberg-news