POC

POC 1:

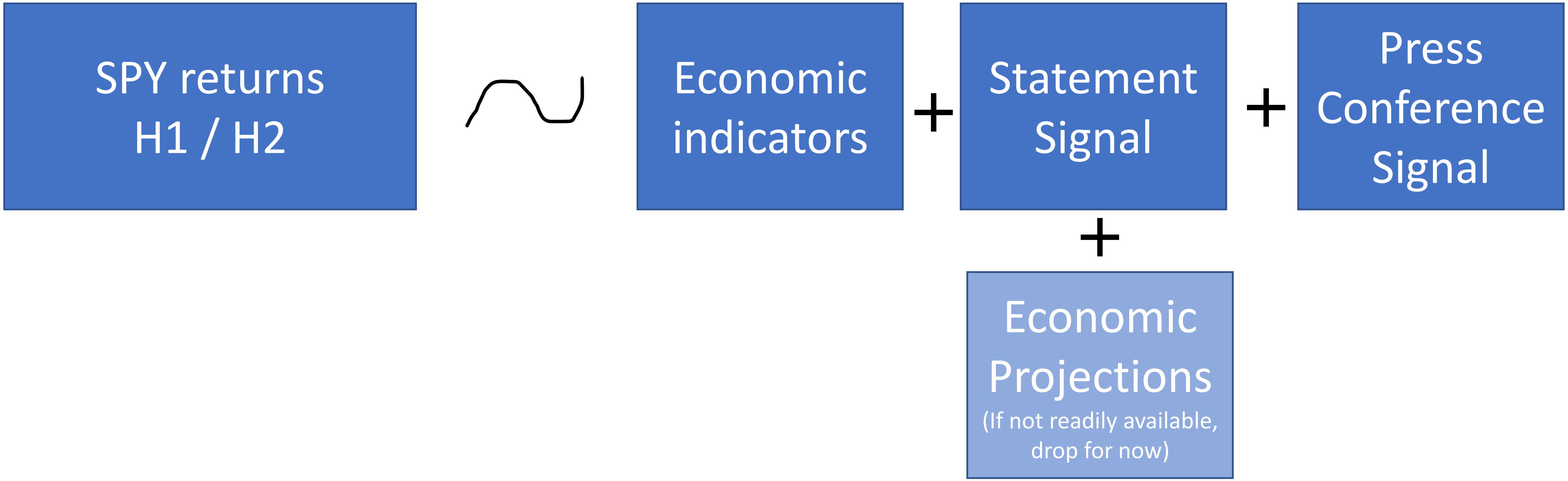
- What causes the volatility in the market? Is it just economic condition?
- What effect does statement have on this volatility?
- What effect does press conf have on this volatility?

POC 2:

- Can LLM generate the same signal as in the original press conference?

Model training

Model


SPY returns H1 / H2
~
Economic indicators
+
Statement Signal
+
Press Conference Signal
+
Economic Projections
(If not readily available, drop for now)

Press conf as Input to NeurNet

- Breakdown Press conf into Opening statement, and individual Q/A pairs for the same DATE
- Final Signal from press conf will be (for each DATE):
  - Opening statement 1. polarity, 2. complexity, and 3. clarity
  - Average question 1. polarity and 2. complexity
  - Average answer:
    1. Polarity (finBERT and other classification models)
    2. Word complexity (CEFR?)
    3. Sentence complexity
    4. Sentence clarity and precision
    5. Answer length
  - Total Conference duration (does it change?)

Statements as Input to NeurNet

- Polarity (finBERT and other classification models)
- Word complexity (CEFR?)
- Document length

Data collection:

- Consensus data: Refinitiv
- Asset prices
  - Price / Returns and Trade volume
  - S&P 500 (Or SPY ETF), 1yr, 2yr, 5yr Treasury yield
  - **1 or 5 minute frequency**.

Model Prediction

LLM Fine tuning

Task 1:

Generate a template of questions given statement, EconInd as context.

Task 2:

Generate answers to the questions given the statement, EconInd as context.

Task 3 (could be FUTURE extension):

Generate a statement given the economic indicators and market projections

Data Augmentation

- Create more Q/A pair using GPT
- Simulate
    - additional answers to the same question
    - additional questions to the same answer