

# Press Release Sentiments for Price Prediction

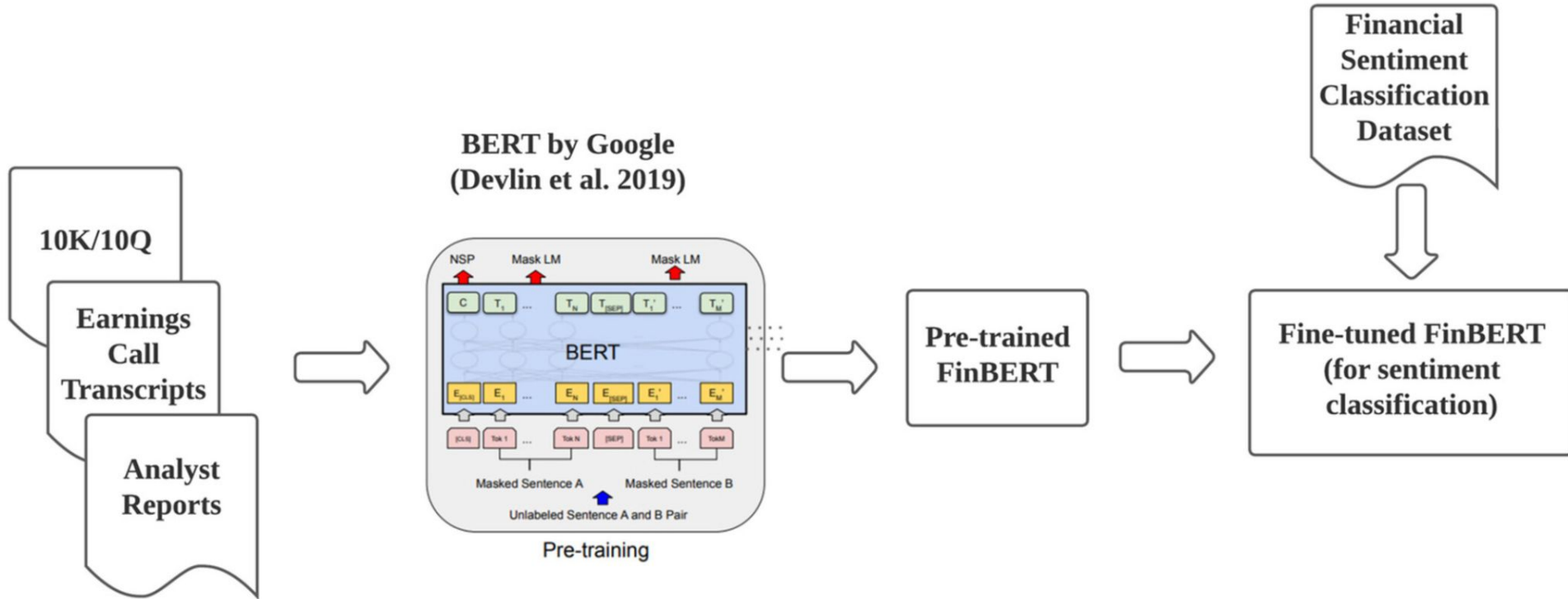
Huseyin Mert Ezer

# What is the research topic and design idea?

- Could press releases have an influence on stock prices?

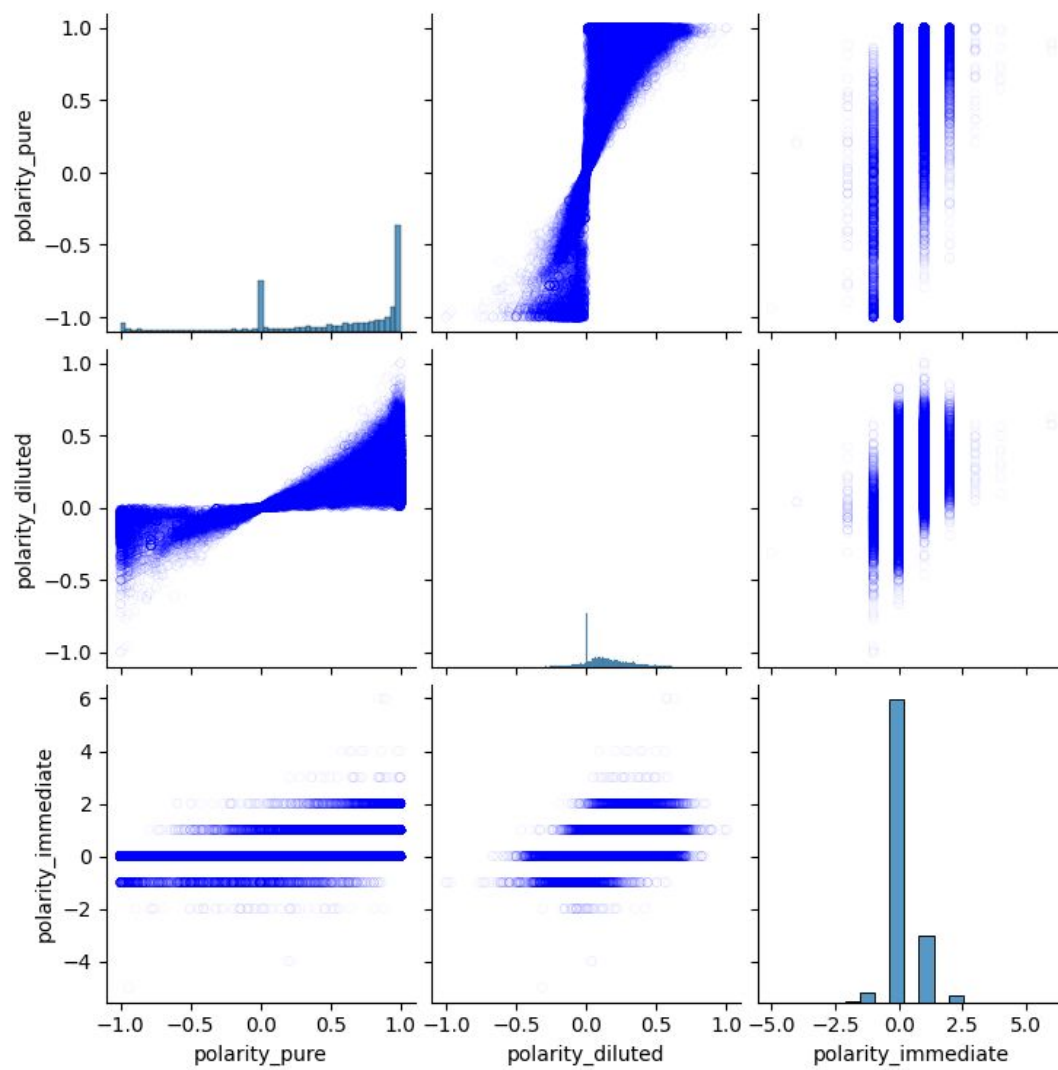
-> Derive sentiment scores from the text of the press releases, and build a model for predicting future movement of the stock price.

# FinBERT



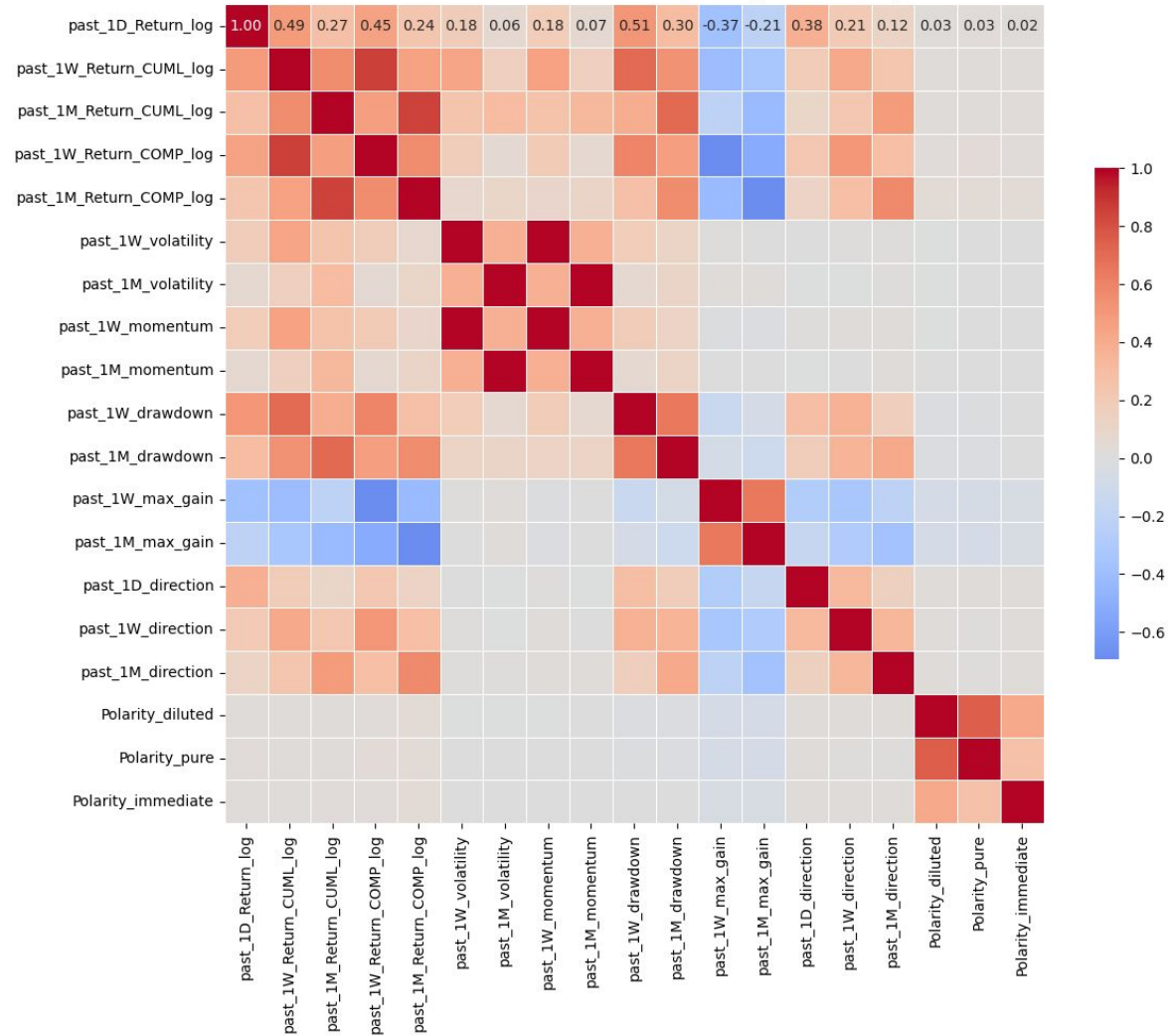
# Sentiment score extraction with FinBERT

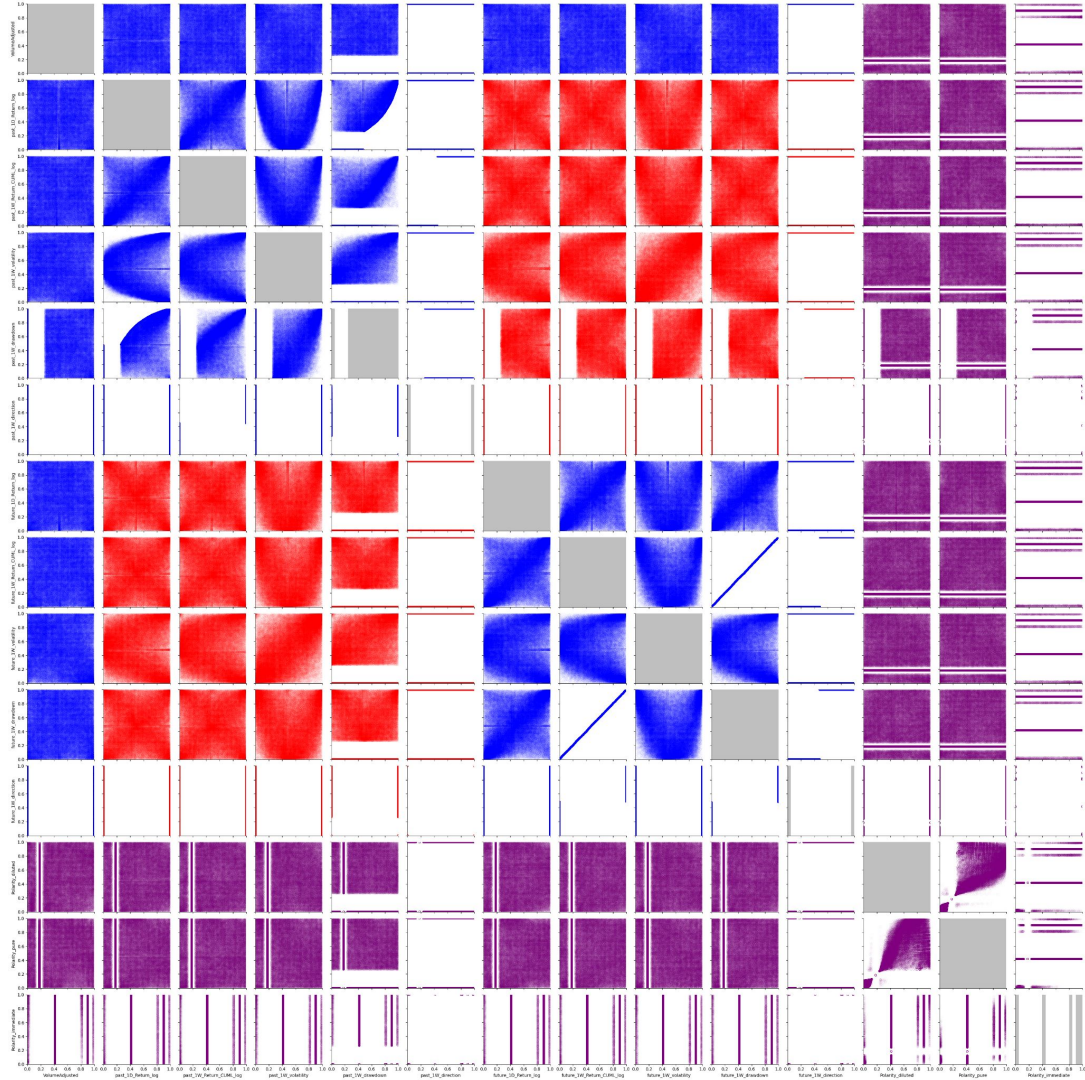
- Preprocess and format the sentences of the press release.
- Run sentiment score extraction with FinBERT at three levels:
  - `polarity\_pure` contains the cumulative value of the sentences that are \*polar\*.
  - `polarity\_diluted` dilutes the scores of the polar sentences proportional to the text length.
  - `polarity\_immediate` sums up the polarity direction of the polar sentences.



# Feature Engineering

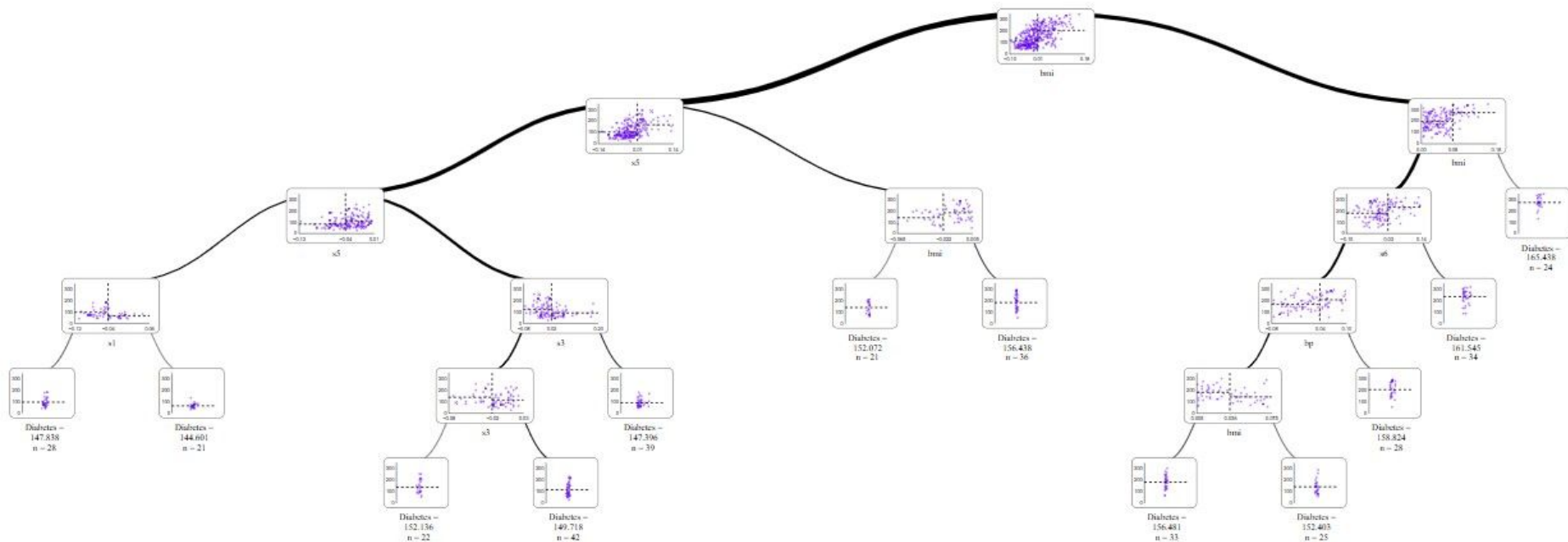
### Correlation Matrix of Past Features (Heatmap Checkerboard)



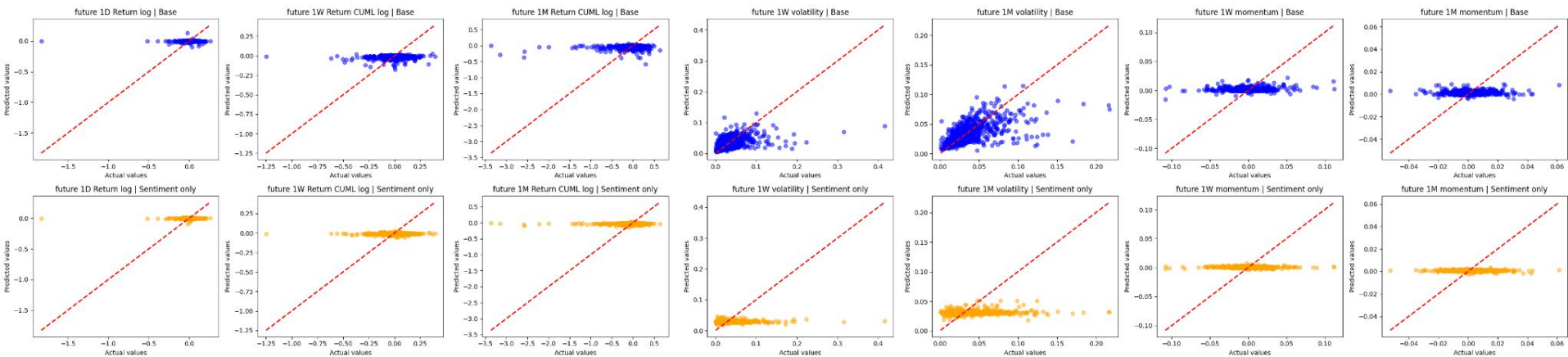




# Model Selection: LightGBM. But why?



# RESULTS



Model Type	MAE		MSE	
	base	sentiment-only	base	sentiment-only
Target				
future_1D_Return_log	0.030978	0.031106	0.006161	0.006097
future_1M_Return_CUML_log	0.143166	0.143970	0.085261	0.087498
future_1M_momentum	0.005990	0.006032	0.000083	0.000084
future_1M_volatility	0.012190	0.017815	0.000412	0.000678
future_1W_Return_CUML_log	0.063410	0.063740	0.010749	0.010870
future_1W_momentum	0.011884	0.012050	0.000349	0.000361
future_1W_volatility	0.014816	0.019513	0.000741	0.001014

Model Type	RMSE		R2	
	base	sentiment-only	base	sentiment-only
Target				
future_1D_Return_log	0.078492	0.078082	-0.011291	-0.000736
future_1M_Return_CUML_log	0.291995	0.295800	0.026188	0.000643
future_1M_momentum	0.009113	0.009145	-0.004495	-0.011694
future_1M_volatility	0.020307	0.026033	0.398673	0.011766
future_1W_Return_CUML_log	0.103676	0.104258	0.007739	-0.003442
future_1W_momentum	0.018689	0.018999	0.018653	-0.014146
future_1W_volatility	0.027220	0.031844	0.268543	-0.001093

Model Type	MAPE	
	base sentiment-only	
Target		
future_1D_Return_log	1.636231e+11	3.576905e+11
future_1M_Return_CUML_log	1.842182e+12	1.100059e+12
future_1M_momentum	2.592350e+00	7.039294e+00
future_1M_volatility	4.840562e-01	9.120990e-01
future_1W_Return_CUML_log	5.908360e+11	6.049956e+11
future_1W_momentum	1.635182e+00	1.563163e+00
future_1W_volatility	7.600170e-01	1.237338e+00

How to improve?