

Prediksi Harga Bitcoin Menggunakan Transformer

Mengintegrasikan Harga, Sentimen, Tren, dan Volume dalam Analisis Deret Waktu Multivariat



Gilang Islamay Putra Djuharis (Soon to be S.Mat)
11190940000055

Ide Utama

01

**Analisis
Sentimen**



02

**Peramalan
Deret Waktu**

Penelitian Terdahulu

Q1

Forecasting

On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles

Big Data and Cognitive Computing

Predicting the Price of Bitcoin Using Sentiment-Enriched Time Series Forecasting

Q1

Q1

Journal of Big Data

Multivariate cryptocurrency prediction: comparative analysis of three recurrent neural networks approaches

Neural Computing and Applications

Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach

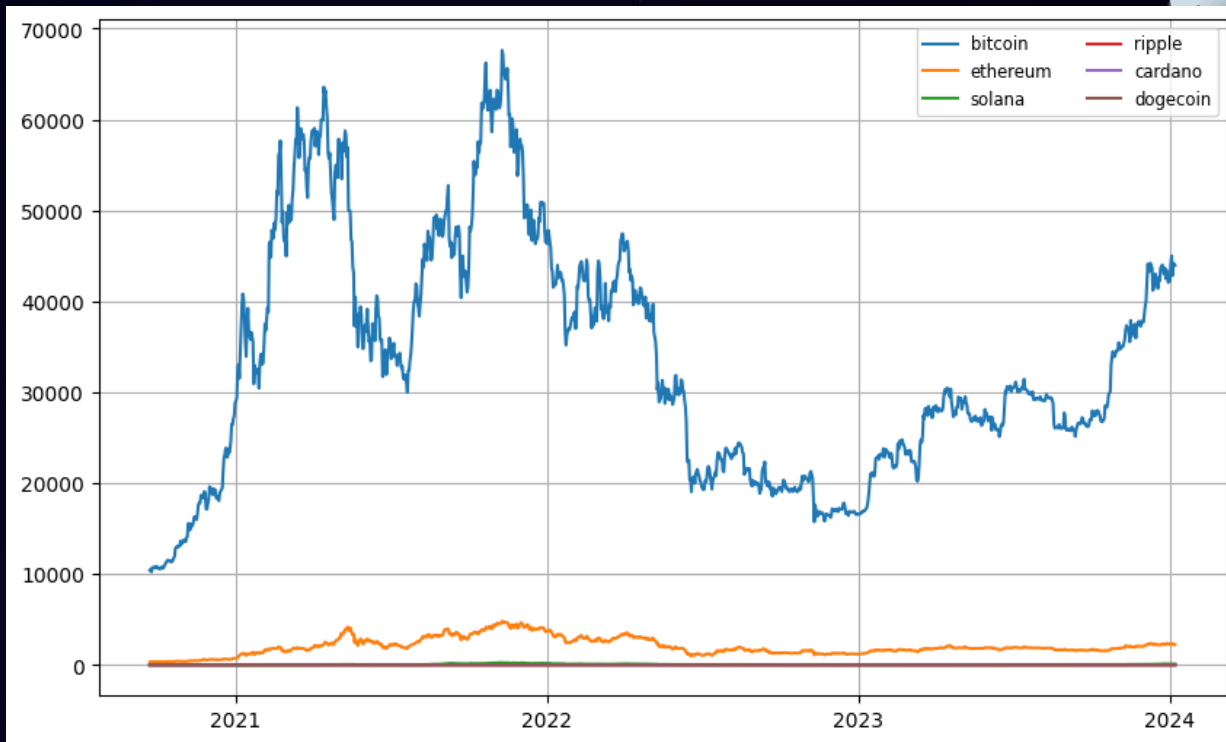
Q1

Kenapa Mata Uang Kripto?

Bitcoin | Ethereum | SOLANA | XRP | Cardano | Avalanche | Dogecoin

Harga

This Price don't Lie



(API coingecko.com)

Popularitas



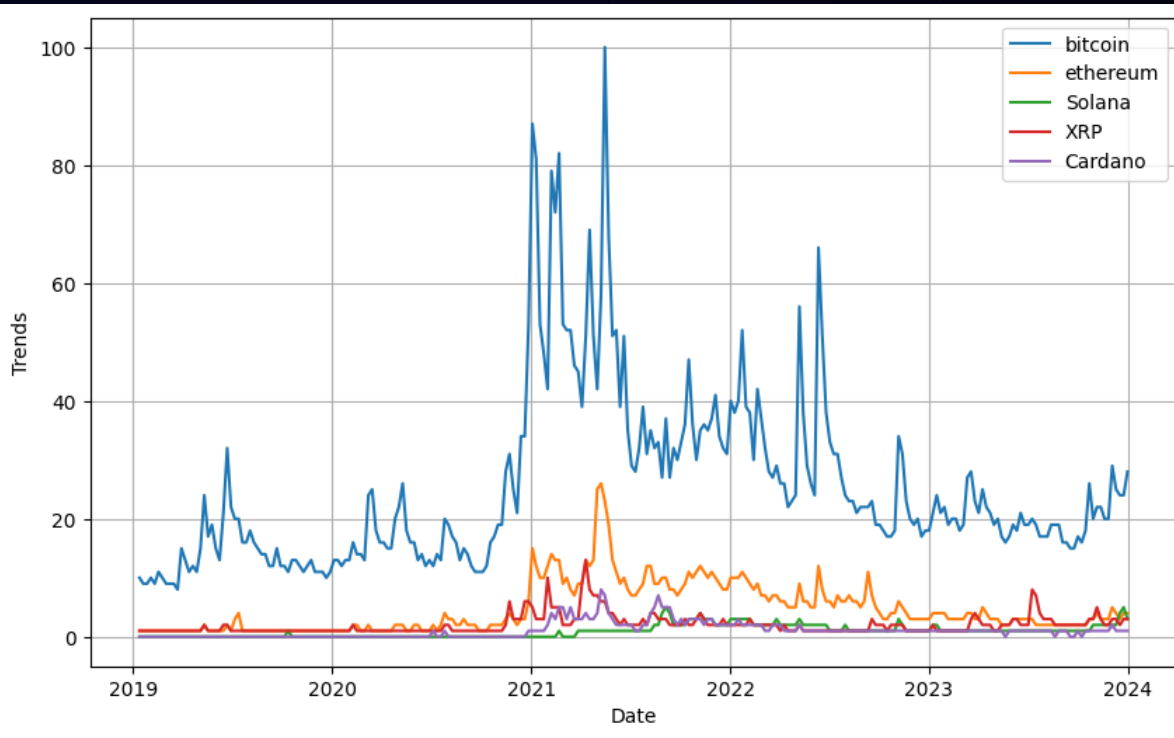
WILLIAM SUBERG

**Bitcoin Price Reaches
Highest Since 2017**

(cointelegraph.com, 2021)

**Bitcoin
Time High
\$69,000
Factors**

(forbes.com, 2021)



(trends.google.com)

**Year as
Scandals**

% this year

h of almost \$69,000

(bloomberg.com, 2023)

*This Time
le*

(nytimes.com, 2020)



Bagaimana?

By paying attention!



A vertical timeline diagram on a dark blue background with star-like bokeh effects. It consists of four white circles arranged vertically, connected by downward-pointing arrows. Each circle contains a year, and to its right is the name of a model or paper. The first circle is at the top, followed by the second, then the third, and finally the fourth at the bottom. The years are 2017, 2018, 2019, and 2019 respectively. The models are 'Attention Is All You Need', 'BERT', 'roBERTa', and 'Temporal Fusion Transformer'.

2017

Attention Is All
You Need

2018

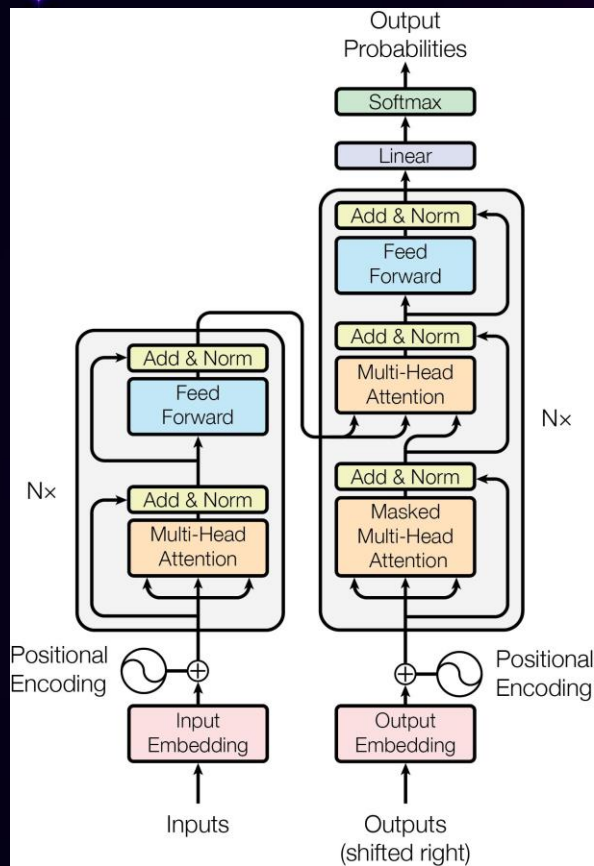
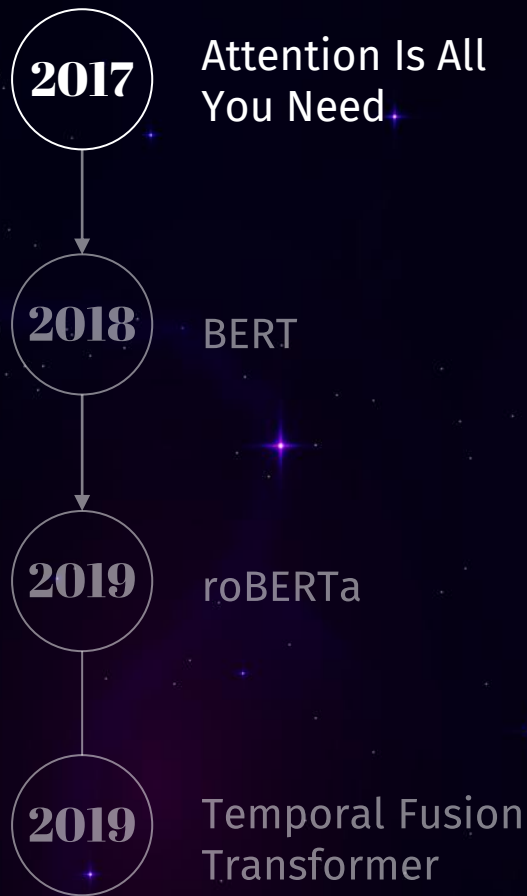
BERT

2019

roBERTa

2019

Temporal Fusion
Transformer



(Attention Is All You Need, 2017)

2017

Attention Is All
You Need

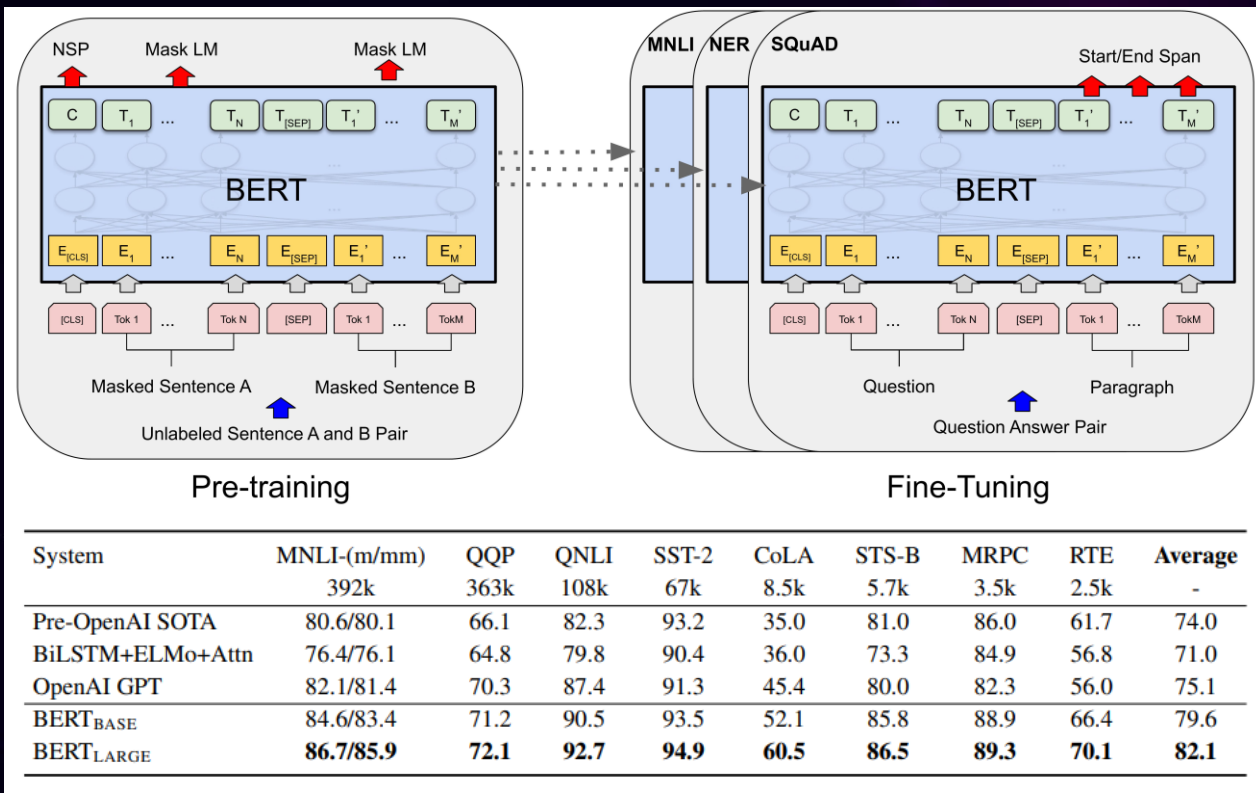
2018

BERT

2019

roBERTa

2019

Temporal Fusion
Transformer(BERT: Pre-training of Deep Bidirectional
Transformers for Language Understanding, 2018)

2017

Attention Is All
You Need

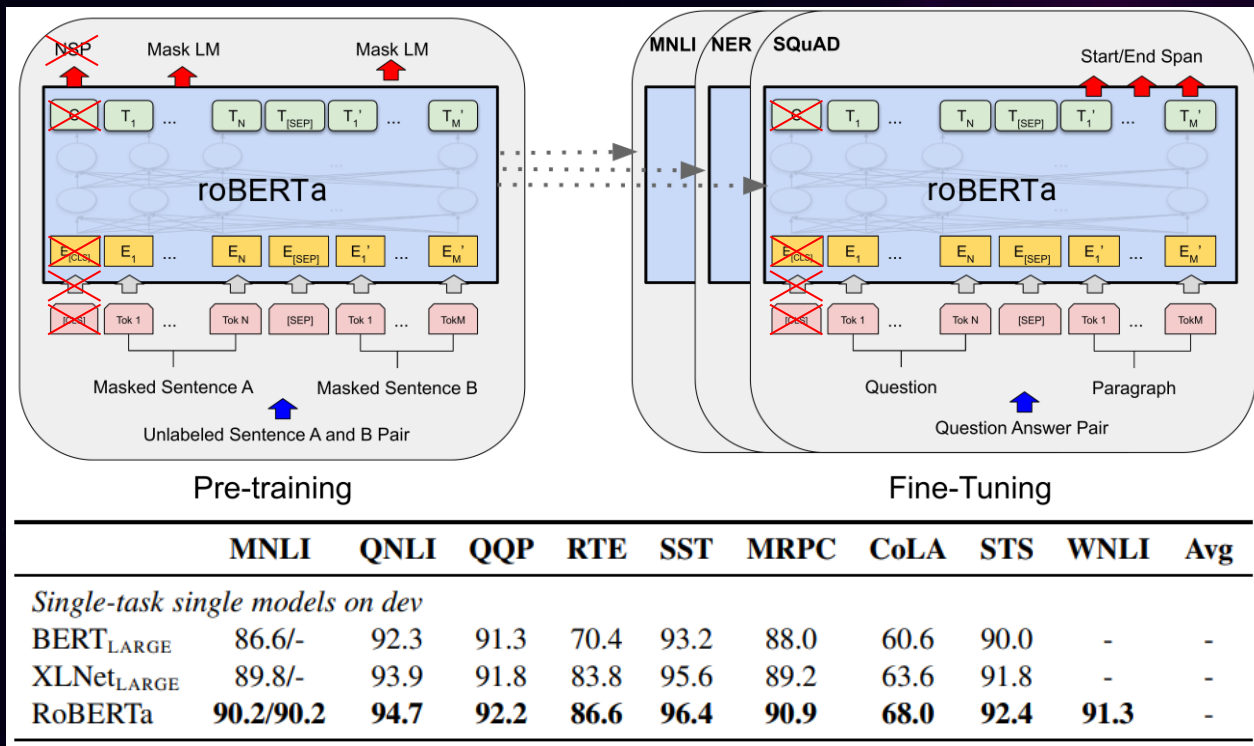
2018

BERT

2019

roBERTa

2019

Temporal Fusion
Transformer(RoBERTa: A Robustly Optimized BERT
Pretraining Approach, 2019)

2017

Attention Is All You Need

2018

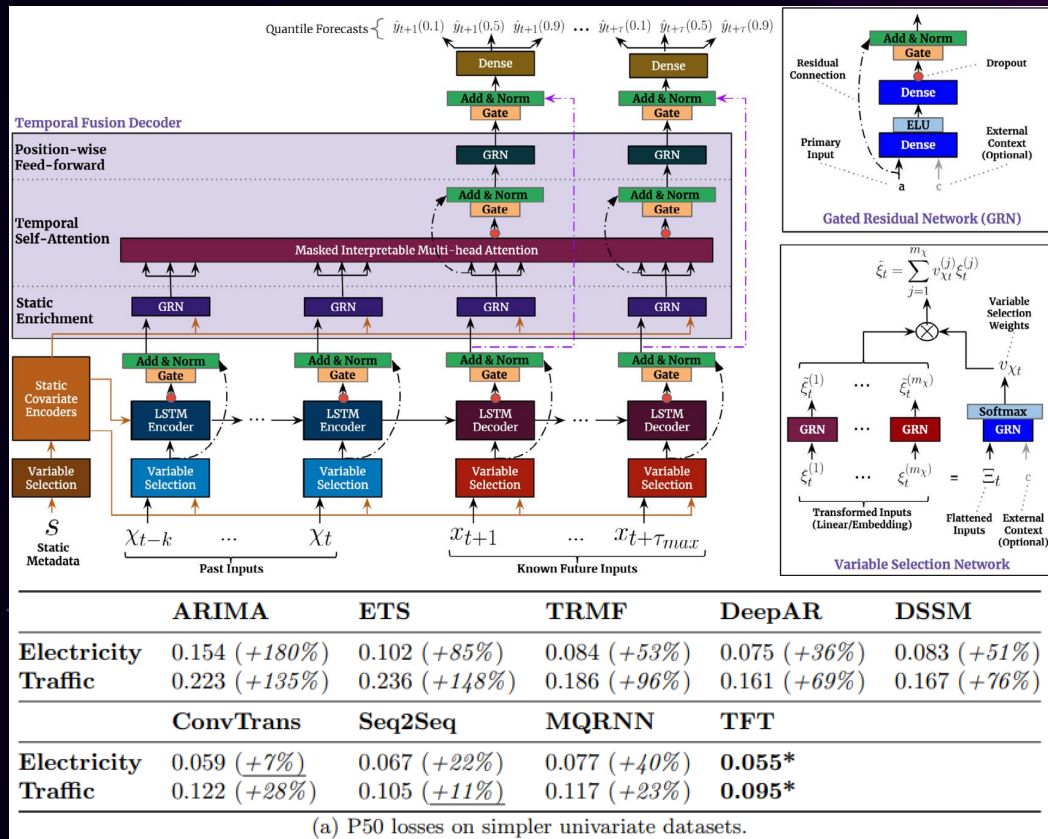
BERT

2019

roBERTa

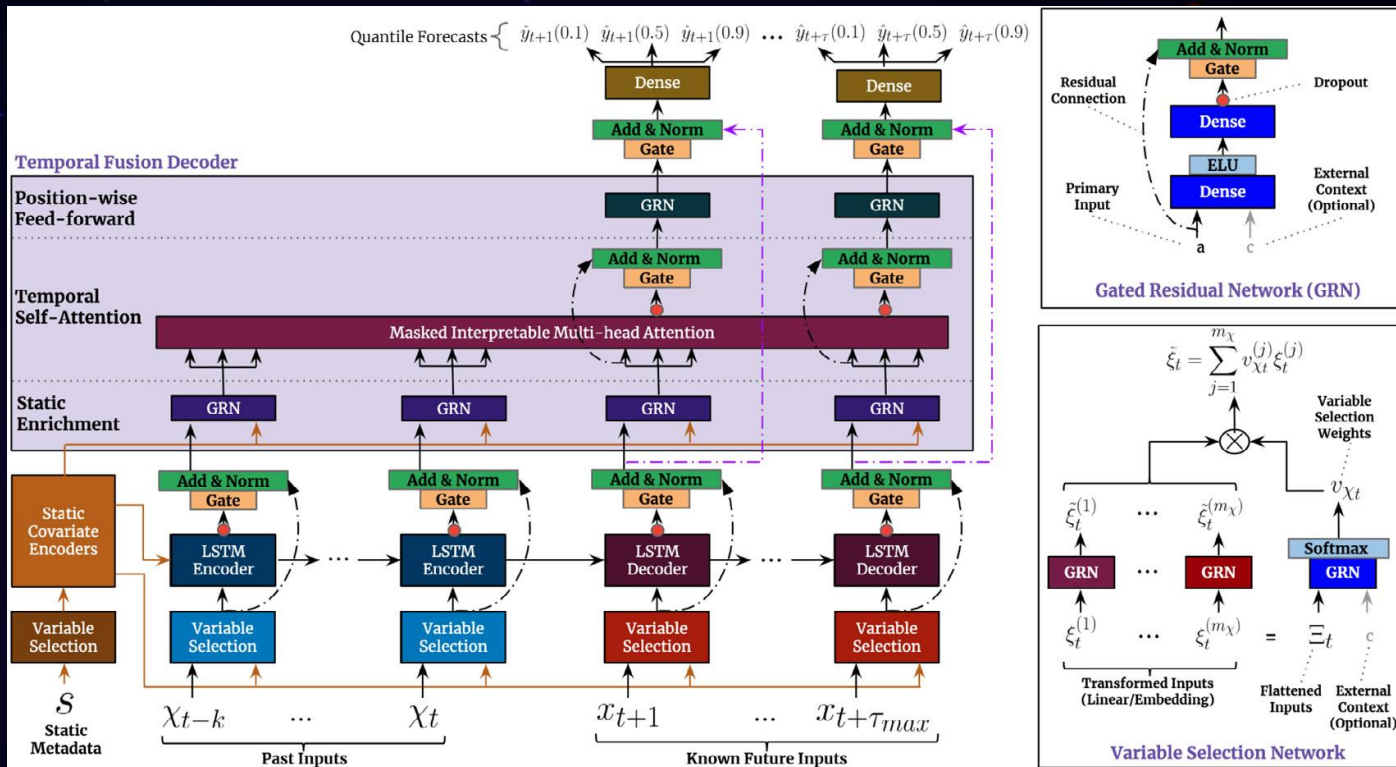
2019

Temporal Fusion Transformer



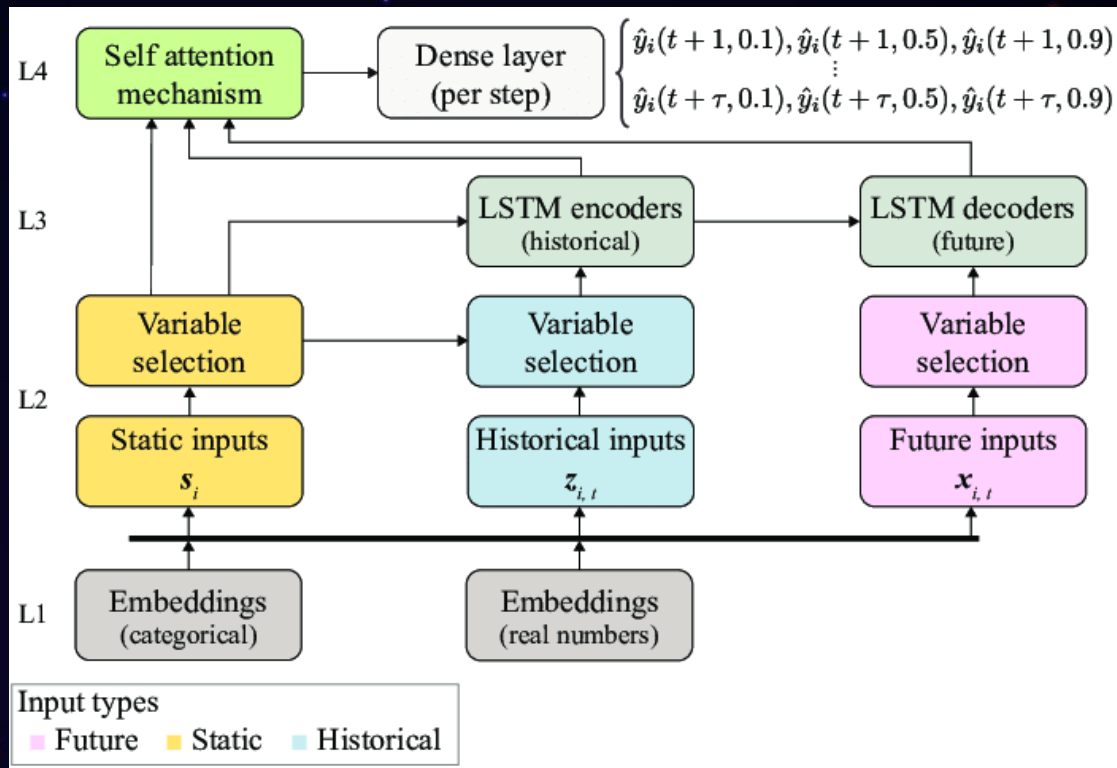
(Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting, 2019)

Temporal Fusion Transformers



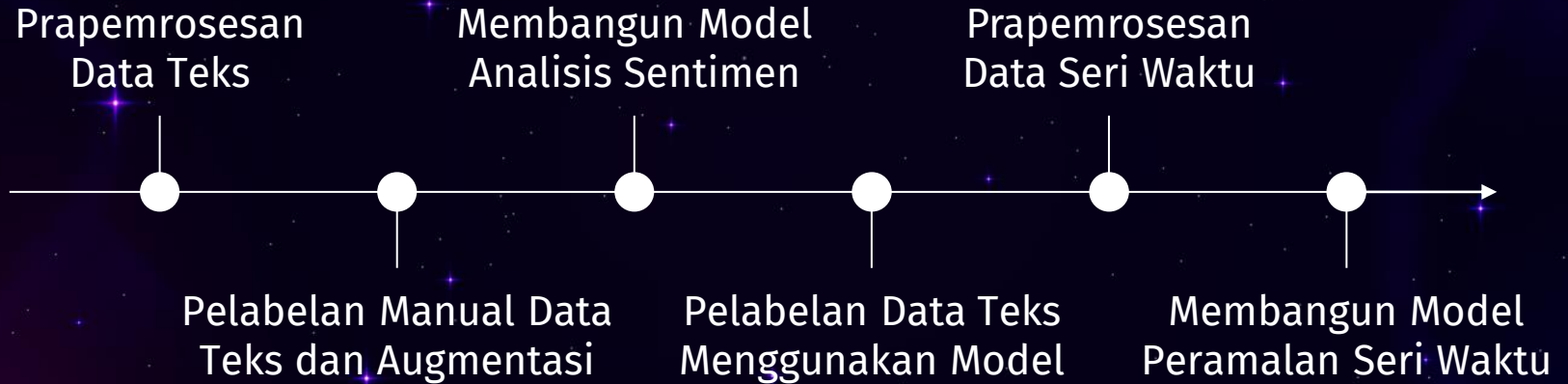
(Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting, 2019)

Temporal Fusion Transformers



(Extreme Precipitation Seasonal Forecast
Using a Transformer Neural Network, 2021)

Jalur Penelitian



01 Analisis Sentimen

Tujuan

Membangun model klasifikasi sentimen terhadap topik Bitcoin

Prapemrosesan Data Teks

01

Penghapusan URLS
dan Entitas HTML

02

Penyelarasan
Huruf Kecil

03

Tokenisasi

04

Penghapusan Tanda
Baca dan Kata - Kata
Penghubung



05

Penghapusan Karakter
Spesial dan Angka

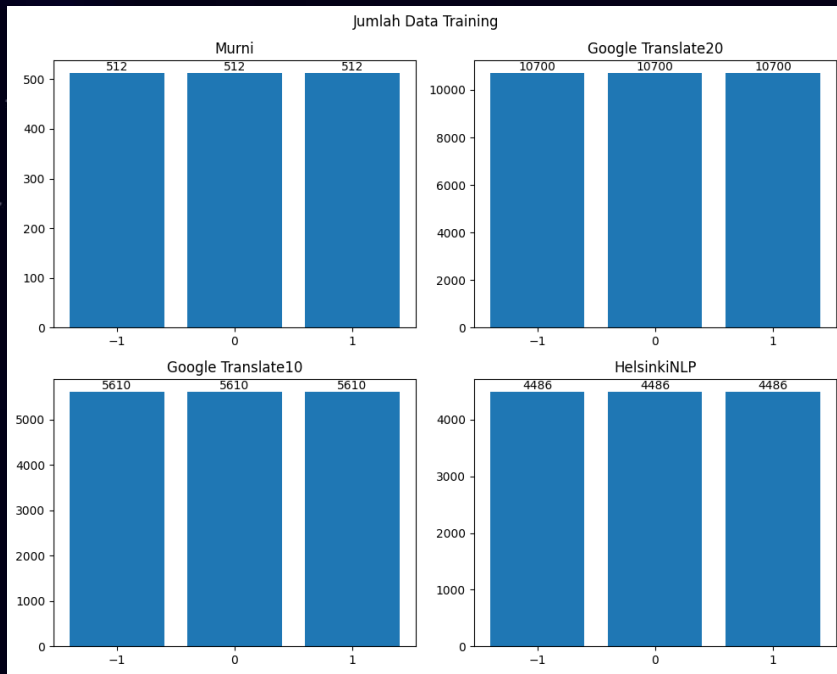
06

Lematisasi

Penghapusan Tweet yang bersifat Spam

date	Text
2021-07-03 20:53:51	<p>The \$BTC price is at \$34655.23 right now.</p> <ul style="list-style-type: none">● Compared to the last tweet, the price has dropped by \$72.77 (-0.21%).● In the last 24 hours the price has increased by \$1488.36 (4.49%). <p>#Bitcoin #BTC https://t.co/5wDhcU31UB</p>
2021-10-19 09:16:20	<p>🗉 🏠 Bitcoin Whale Alert: [TX: 6e2f93751abd08b63306d5351da016187927ac712f442d6709d23d10b01e83c8]-[ADDR: 1DuhtLa8TtCC547WfSNLHZvT91b9PQQeDD]-[#BTC: 16.68212472]-[BLOCK_DATE: 2021-10-19 10:37:49] #btc_whale_alert #bitcoin- BTC_Whale_Alert</p>
2021-06-21 07:43:39	<p>WhaleTrades:  💰 \$2,500,000 #bitcoin LONGED @\$32,805.8359 [21/06/21 07:34:55]</p> <p> BitMEX \$XBTUSD</p> <p>💬 I'd take a fast nickel over a slow dime - buyerofblood</p>
2022-04-16 15:59:12	<p>💰 Market Cap. Swap 💰 What would one #Ecash \$XEC cost if it had the market capitalization of #Bitcoin \$BTC? One #XEC would be worth \$0.0401409. 🤖 🐦 ✍️ 👁 More: https://t.co/fSCkKGmhckK... https://t.co/rmygNBMIhr</p>

Pelabelan Data Teks dan Augmentasi



**Google
Translate20**

Top 20 Skor BLEU
Tertinggi di Google
Translate

**Google
Translate10**

Top 10 Skor BLEU
Tertinggi di Google
Translate

HelsinkiNLP

Model Translasi dari
Universitas Helsinki

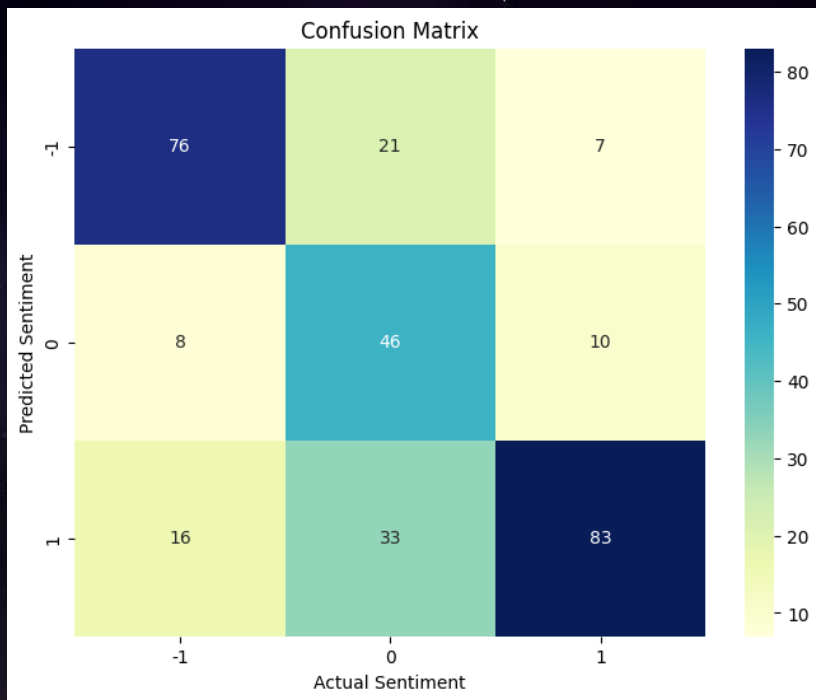
Membangun Model Analisis Sentimen

Model	Acc Negative Sentiment	Acc Neutral Sentiment	Acc Positive Sentiment	Overall Accuracy
twitter-roberta-base-sentiment-latest	0,59	0,72	0,27	0,526666667
bertweet-base-sentiment-analysis	0,53	0,75	0,24	0,506666667
finbert	0,05	0,06	0,91	0,34
twitter-xlm-roberta-base-sentiment	0,52	0,7	0,3	0,506666667
cryptobert	0,37	0,56	0,59	0,506666667
twitter-xlm-roberta-bitcoin-sentiment	0,47	0,51	0,38	0,453333333

Fine Tuned

Model	Sebelum Fine Tuning	Murni	GT20	GT10	HelsinkiNLP	Improvement
bertweet-base-sentiment-analysis	0,5066667	0.626667	0. 666666	0. 666666	0.656667	+31.62%
Cryptobert	0,5066667	0.666666	0.673333	0.676667	0.646667	+33.60%
twitter-roberta-base-sentiment-latest	0,5266667	0.653333	0.65	0.646667	0.64	+24.14%
twitter-xlm-roberta-base-sentiment	0,5066667	0.606667	0.596667	0.616667	0.62	+22.53%

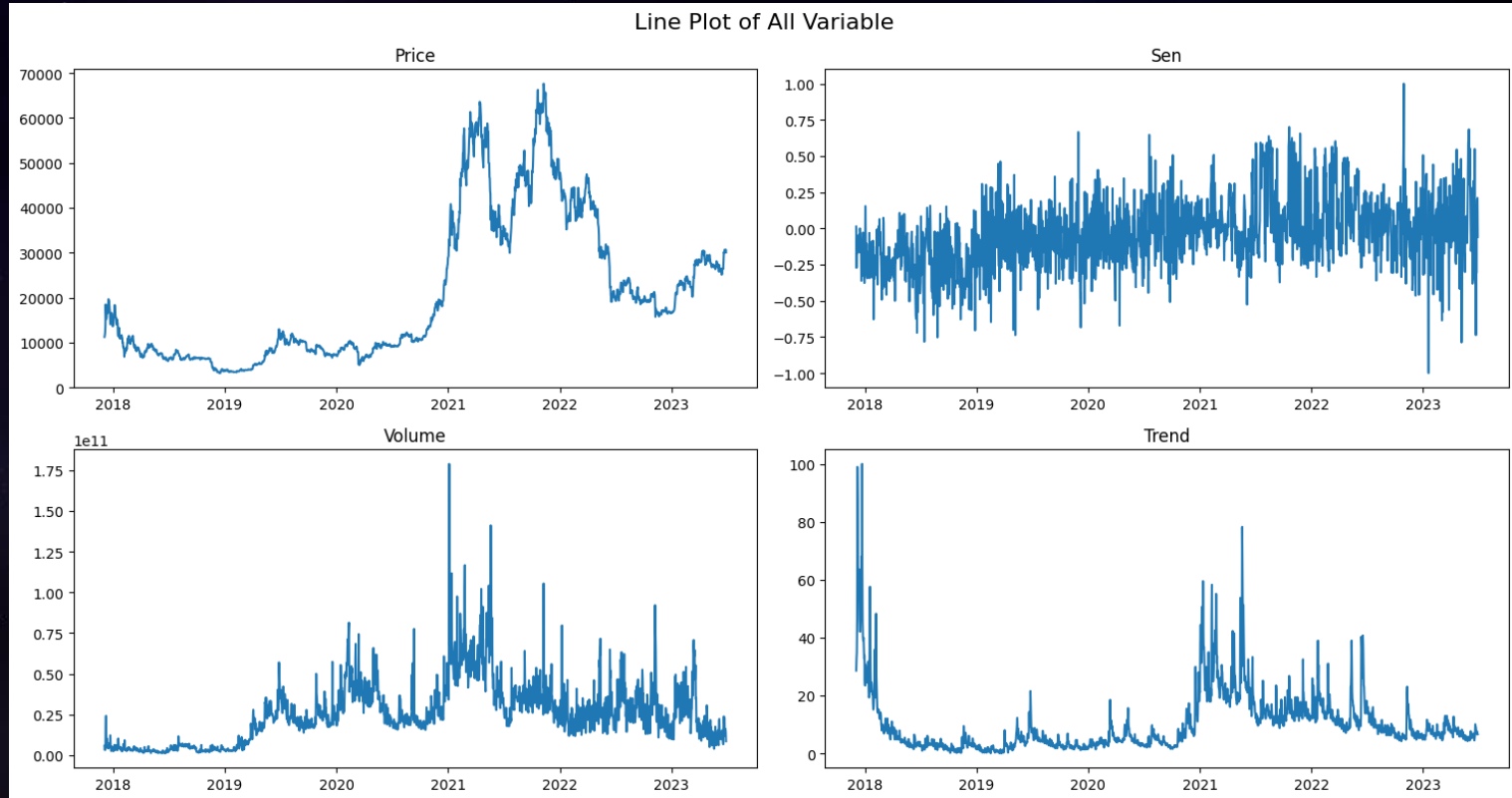
Introducing CryptobertRefined



Sentiment Score	Precision	Recall	F1-Score	n
-1	0,73	0,76	0,75	100
0	0,72	0,48	0,56	100
1	0,63	0,83	0,72	100
Accuracy	0,68			300
Macro Avg	0,69	0,68	0,67	300
Weighted Avg	0,69	0,68	0,67	300



Pelabelan Data Teks menggunakan Model

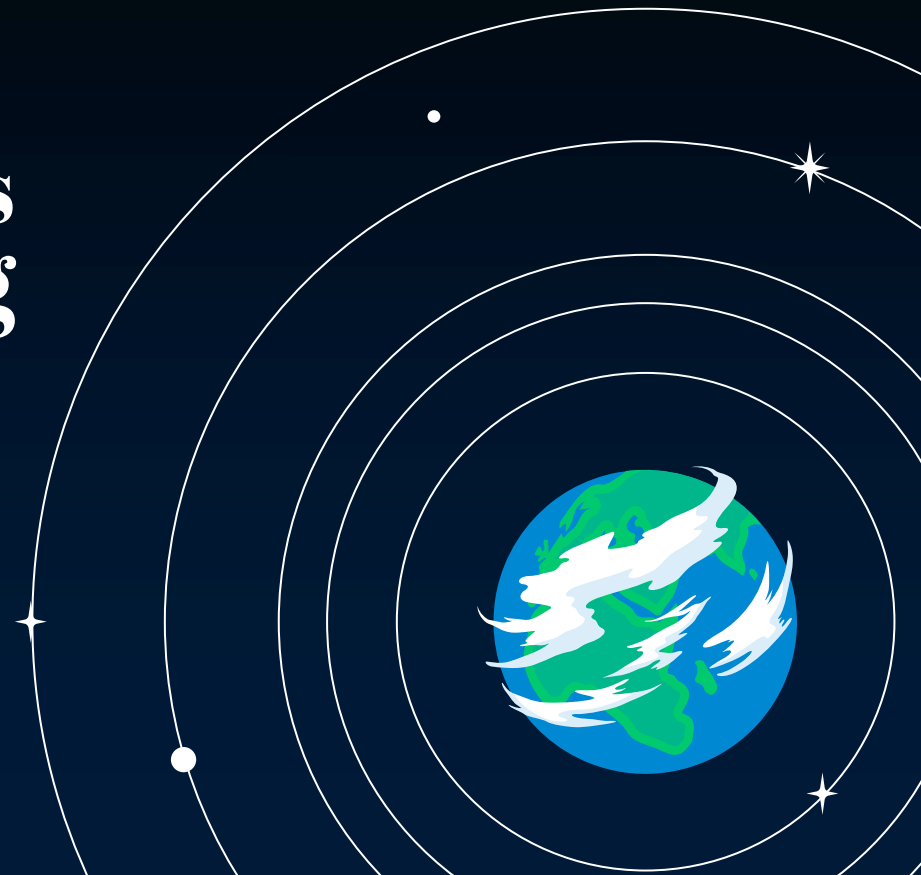


02 Time Series Forecasting

Let's do some Forecasting!

Tujuan

Membangun model Peramalan
Seri Waktu Harga Bitcoin



Prapemrosesan Data Deret Waktu

01

Outlier Detection

Isolation Forest

02

Outlier Handling

Simple Moving Average

03

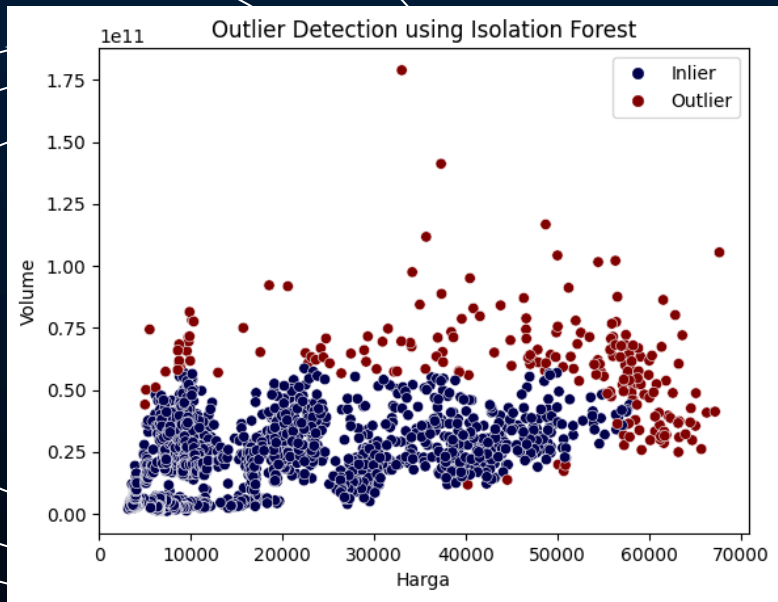
Stationary Test

Augmented Dickey-Fuller Test

04

Transformation

Differencing Transformation



Outlier Detection dan Handling

Terdapat 204 outlier!

Stationary Test

```
from statsmodels.tsa.stattools import adfuller
```

```
#perform augmented Dickey-Fuller test
```

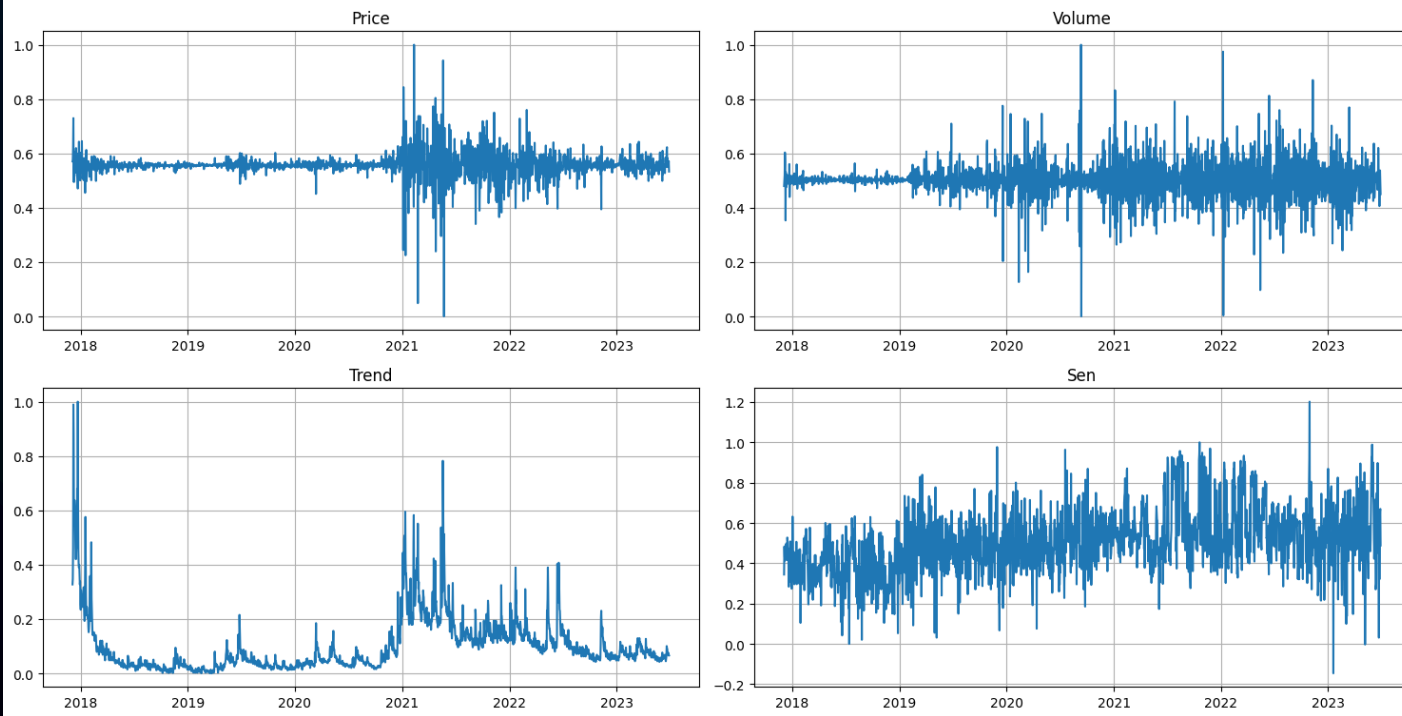
```
print(adfuller(df['price']))  
print(adfuller(df['volume']))  
print(adfuller(df['sen']))  
print(adfuller(df['trend']))
```

```
(-1.3685338553623199, 0.5972650025176758, 21, 2013, {'1%': -3.4336026867364144, '5%': -2.8629768669694458, '10%': -2.5675349256978177}, 34089.00522190065  
4)  
(-2.4631832221188876, 0.12469110322237731, 26, 2008, {'1%': -3.433610796445261, '5%': -2.8629804474582348, '10%': -2.5675368321236967}, 96542.4693010425)  
(-6.099371696292173, 9.9206487232894e-08, 13, 2021, {'1%': -3.4335897947733383, '5%': -2.8629711750645868, '10%': -2.567531895054917}, -825.212502464707  
5)  
(-4.561227032801475, 0.00015217925898935135, 24, 2010, {'1%': -3.4336075477138945, '5%': -2.862979013123536, '10%': -2.567536068414148}, 10240.2990682240  
9)
```

Differencing Transformation

The First Order! With some Spice

Line Plot of All Variables



Stationary Test

```
print(adfuller(df['price']))  
print(adfuller(df['volume']))
```

```
(-7.954355277425575, 3.06555064072429e-12, 26, 2007, {'1%': -3.4336124232421144, '5%': -2.8629811656983755, '10%':  
8})  
(-13.11958536363419, 1.5732211454090421e-24, 26, 2007, {'1%': -3.4336124232421144, '5%': -2.8629811656983755, '10%':  
871})
```

The background is a dark navy blue space-themed illustration. It features several stylized galaxies in shades of teal, green, and purple. White stars of various sizes are scattered throughout. In the top-left and bottom-right corners, there are sets of concentric white circles, resembling orbits or signal waves, with small white dots at their centers. The word "FORECASTING!" is centered in a large, white, serif font.

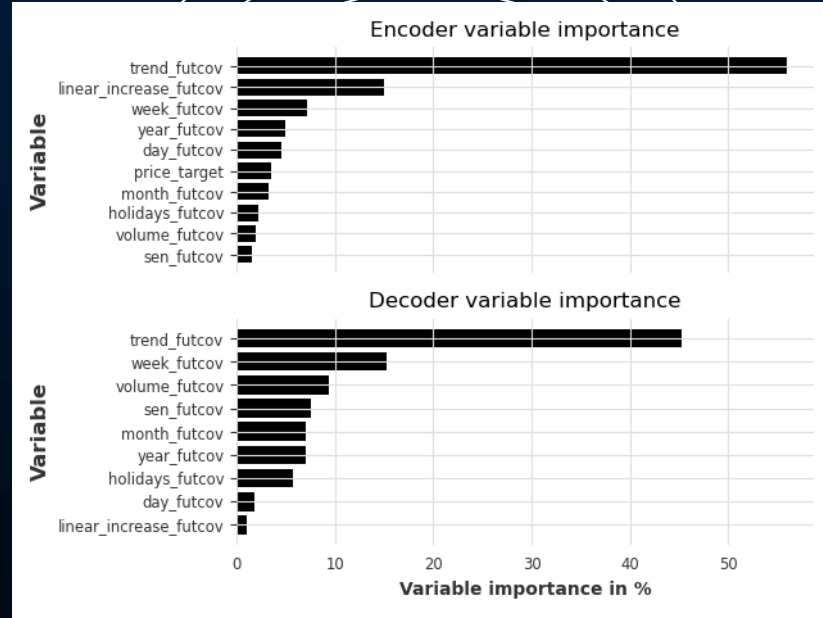
FORECASTING!

Model Peramalan Seri Waktu Harga Bitcoin

Model	Moving Average			Box Plot			Pelatihan (s)
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	
TFT	0,0261	0,0177	3,0973	0,0463	0,0312	6,7257	166,14
TFT _{w/o}	0,0263	0,0176	3,0737	0,0479	0,0316	6,9818	
TCN	0,0263	0,0175	3,0627	0,0477	0,0322	7,0084	46,94
TCN _{w/o}	0,0262	0,0172	3,0166	0,048	0,0318	7,0042	
GRU	0,0262	0,0175	3,0764	0,0475	0,0316	6,9442	77,98
GRU _{w/o}	0,0263	0,0175	3,0678	0,0479	0,0318	7,025	
LSTM	0,0262	0,0174	3,0532	0,0477	0,0314	6,9306	74,05
LSTM _{w/o}	0,0264	0,0174	3,0469	0,0479	0,0317	7,0063	



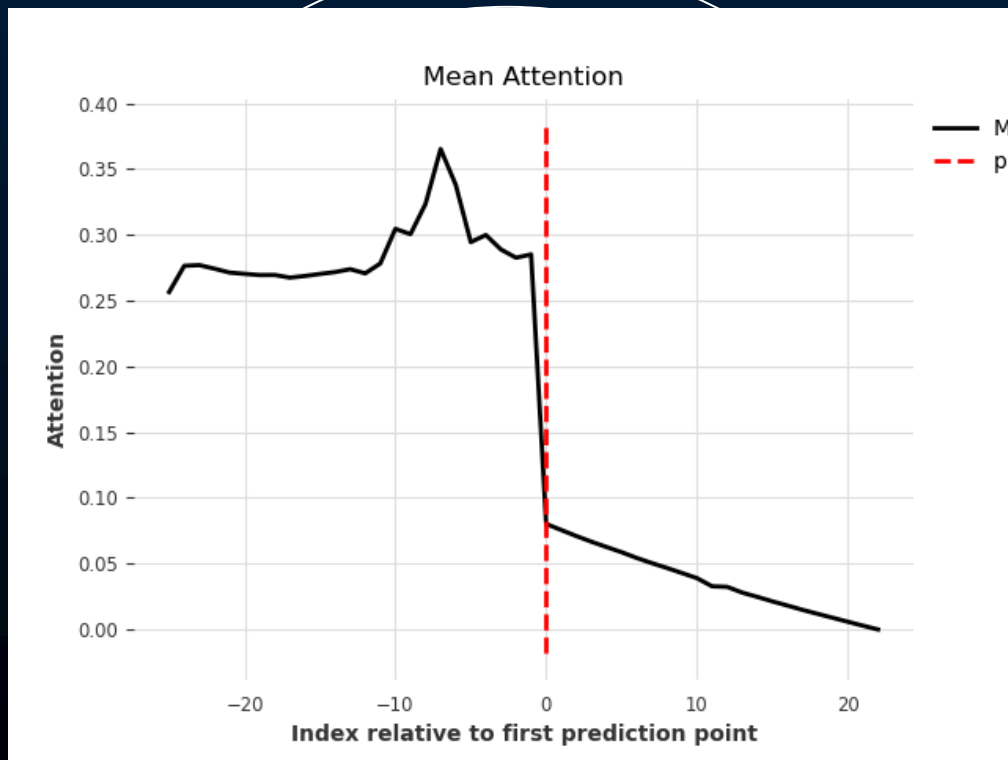
EXPLAINABILITY!



Variable Importance

Type	Encoder	Decoder
sen_futcov	0,02	0,08
volume_futcov	0,02	0,09
holidays_futcov	0,02	0,06
month_futcov	0,03	0,07
price_target	0,04	-
day_futcov	0,05	0,02
year_futcov	0,05	0,07
week_futcov	0,07	0,15
linear_increase_futcov	0,15	0,01
tren_futcov	0,56	0,45

Attention Horizon

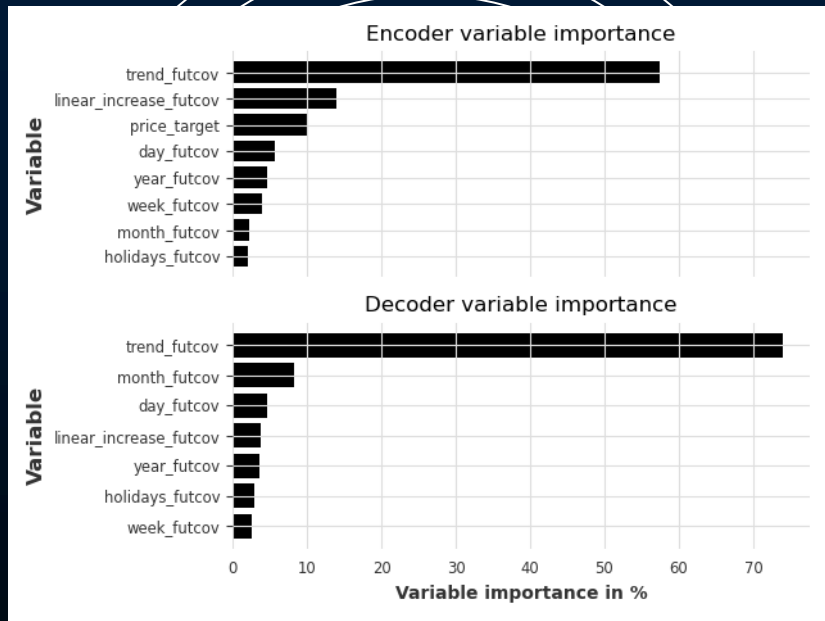


The background is a dark navy blue space-themed illustration. It features several stylized galaxies in shades of teal, green, and purple. White stars of various sizes are scattered throughout. In the top-left and bottom-right corners, there are sets of concentric white circles, some with a star at their center, resembling orbital paths or celestial bodies. The word "EXPERIMENT!" is centered in a large, white, serif font.

EXPERIMENT!

Model Peramalan Seri Waktu Harga Bitcoin

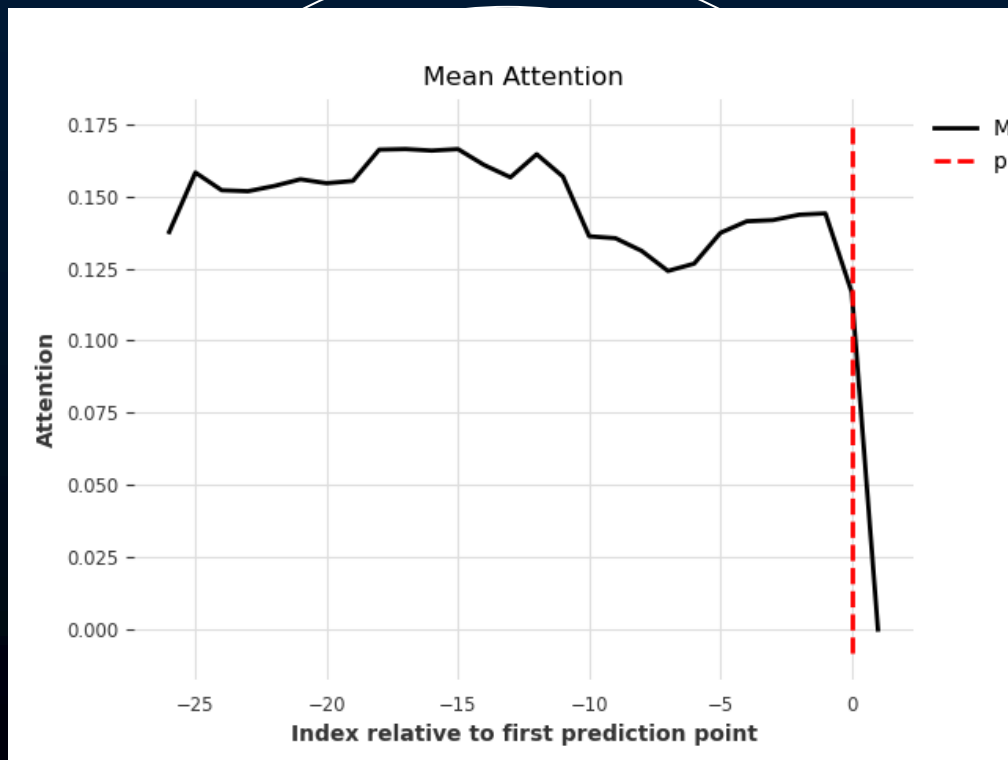
Model	RMSE	MAE	MAPE
TFT (Dengan semua variable)	0,0261	0,0177	3,097
TFT (Tanpa Sentiment dan Volume)	0,0258	0,0173	3,033



Variable Importance

Type	Encoder	Decoder
holidays_futcov	0,02	0,03
month_futcov	0,02	0,08
week_futcov	0,04	0,03
year_futcov	0,05	0,04
day_futcov	0,06	0,05
price_target	0,1	-
linear_increase_futcov	0,14	0,04
tren_futcov	0,58	0,74

Attention Horizon



Kesimpulan

Augmentasi Data

Back-Translation dan
Easy Data Augmentation

Explainability

Kepentingan Encoder
Decoder dan Attention



Data Quality

Google Translate dan
HelsinkiNLP

Variable

Sentimen, Volume, Trend,
dan variabel statis

Forecasting Model

TFT, LSTM, GRU, & TCN



Thank You

“Any Sufficiently Advanced Technology is Indistinguishable from Magic”
- Arthur C. Clarke



Thank You

“Barang siapa yang menjadikan mudah urusan orang lain, pasti Allah akan memudahkannya di dunia dan di akhirat.” (HR. Muslim)