# 1 Domain Background

A set of data that is indexed by time is known as a time series. They appear in many different fields, such as statistics, physics, finance, economics, biology, or even business. Because of their wide applicability, it is important to generate accurate forecasts of of these time series data. These forecasts are generated using specific mathematical models or algorithms which are trained on a subset of the past values of a given time series. For the purspose of simplifying future discussions, we will adopt the following notation for a time series $X(t)$ or $X_t$

$$\{X(t); t = 0, 1, ...\} \tag{1}$$

My personal motivation for working on time series is to use it to take advantage of the best exchange rates. As someone who lives abroad in a country that uses a different currency, I often need to transfer money to-and-from my different bank accounts. These transfers are subject to fluctuating currency exchange rates. Without a way to predict what the exchange rate will be for the time of the transfer, I end up loosing money in these tranfers. Therefore, I am interested in developing a way to forecast the exchange rates so that I can minimize the loses during these transfers.

# 2 Problem Statement

# 3 Datasets and Inputs

The dataset that we will use is taken from Kaggle [https://www.kaggle.com/meehau/EURUSD/home] Exchange Rate TWI. May 1970 – Aug 1995. [https://datamarket.com/data/set/22tb/exchange-rate-twi-may-1970-aug-1995!ds=22tbdisplay=line]

Exchange rate of the Australian Dollar [https://datamarket.com/data/set/22wv/exchange-rate-of-australian-dollar-a-for-1-us-dollar-monthly-average-jul-1969-aug-1995!ds=22wvdisplay=line]

We will also use some data sets for benchmarking our methods

Sunspot data [https://datamarket.com/data/set/22wh/wolfer-sunspot-numbers-1770-to-1869!ds=22whdisplay=line] [https://datamarket.com/data/set/22wg/wolfs-sunspot-numbers-1700-1988!ds=22wgdisplay=line]

Monthly U.S air passenger miles January 1960 through December 1977, n=216 https://datamarket.com/data/set/22sj/monthly-us-air-passenger-miles-january-1960-through-december-1977-n216!ds=22sjdisplay=line

Mean daily temperature, Fisher River near Dallas, Jan 01, 1988 to Dec 31, 1991 https://datamarket.com/data/set/235d/mean-daily-temperature-fisher-river-near-dallas-jan-01-1988-to-dec-31-1991!ds=235ddisplay=line

Total annual rainfall (in inches), London, England, 1813 – 1912 https://datamarket.com/data/set/22np/to annual-rainfall-in-inches-london-england-1813-1912!ds=22npdisplay=line

Stock Data IBM [https://datamarket.com/data/set/2321/ibm-common-stock-closing-prices-daily-29th-june-1959-to-30th-june-1960-n255!ds=2321display=line]

Number of earthquakes per year magnitude 7.0 or greater. 1900-1998 https://datamarket.com/data/set/22 of-earthquakes-per-year-magnitude-70-or-greater-1900-1998!ds=22p8display=line

# 4 Solution Statement

To solve the problem

The models that we will use in the problem solution is

# 5 Benchmark Model

For the benchmark model, we will use a simple linear regression model.

# 6 Evaluation Metrics

$$MAE = \frac{1}{N}\sum_{t=0}^{N}|e_t| \tag{2}$$

$$MAPE = \tag{3}$$

$$MSE = \frac{1}{N}\sum_{t=0}^{N}e_t^2 \tag{4}$$

$$RMSE = \tag{5}$$

$$U = \tag{6}$$

# 7 Project Design

Data preprocessing
Splitting the data

The models that will be considered are LSTM (Long-short term memory neural network), ANN (artificial neural network), A generative adversarial network along with linear regression models and ARMA models.

# References


[1] An Introductory Study on Time Series Modeling and Forecasting LAP Lambert Academic Publishing, Germany, 2013 https://arxiv.org/pdf/1302.6613.pdf

[2] Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The LaTeX Companion*. Addison-Wesley, Reading, Massachusetts, 1993.

[3] Albert Einstein. *Zur Elektrodynamik bewegter Körper*. (German) [*On the electrodynamics of moving bodies*]. Annalen der Physik, 322(10):891–921, 1905.

[4] Knuth: Computers and Typesetting,
`http://www-cs-faculty.stanford.edu/~uno/abcde.html`