Financial time series forecasting applying deep learning algorithms

Erik Solís, Sherald Noboa, Erick Cuenca

erik.solis@yachaytech.edu.ec

2021-11-24

1 Introduction
- Sequence Models
- Time Series
- Trading Strategies

2 Materials

3 Methodology

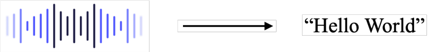
4 Experiments

5 Results

6 Conclusions

Examples of sequence models

- Speech recognition



Examples of sequence models

- Speech recognition

"Hello World"

- Machine translation

Bonjour, comment allez-vous?

¿Buenos días, Cómo está?

Examples of sequence models

- Speech recognition

"Hello World"

- Machine translation

Bonjour, comment allez-vous?

¿Buenos días, Cómo está?

- Sentiment classification

"We didn't like this movie"

Examples of sequence models

- Speech recognition

→ "Hello World"

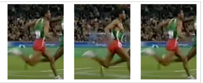
- Machine translation

Bonjour, comment allez-vous? → ¿Buenos días, Cómo está?

- Sentiment classification

"We didn't like this movie" →

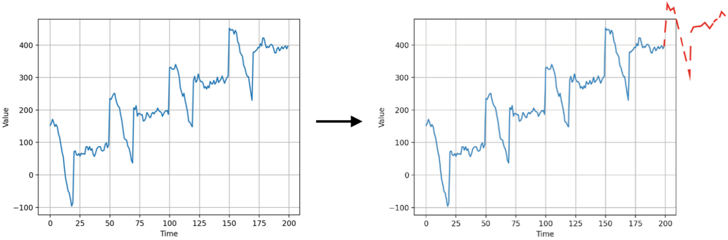
- Video activity recognition



→ Running

Examples of sequence models

- Time series sequence models



1 Introduction
- Sequence Models
- Time Series
- Trading Strategies

2 Materials

3 Methodology

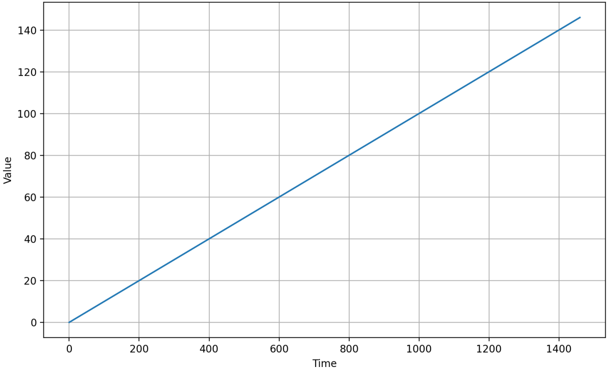
4 Experiments

5 Results

6 Conclusions

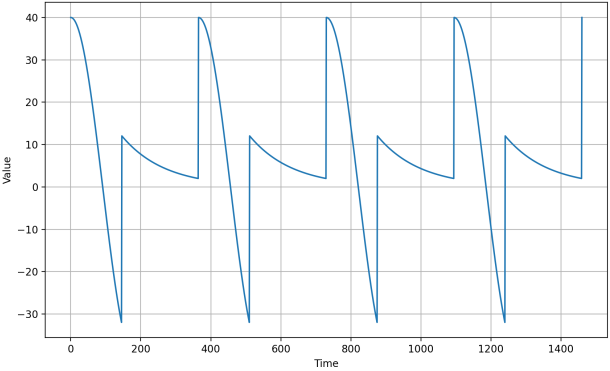
Trend

- Moving in a specific direction



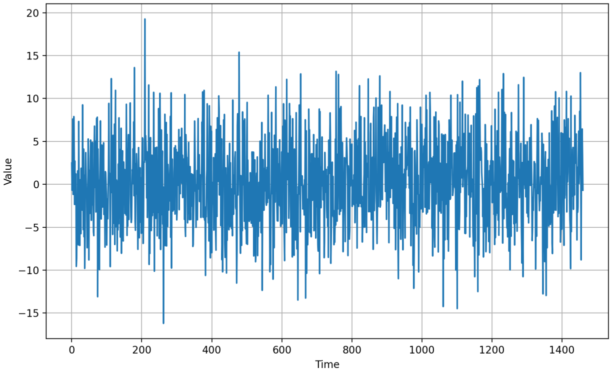
Seasonality

- Patterns repeats in a predictable interval



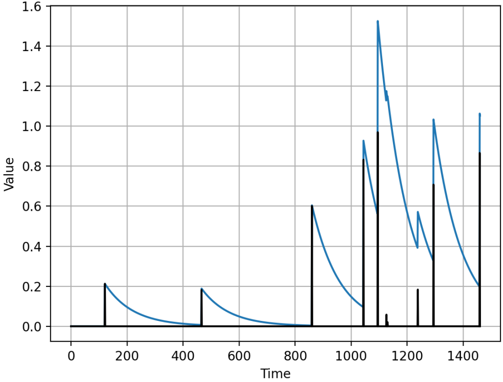
Noise

- Set of random values



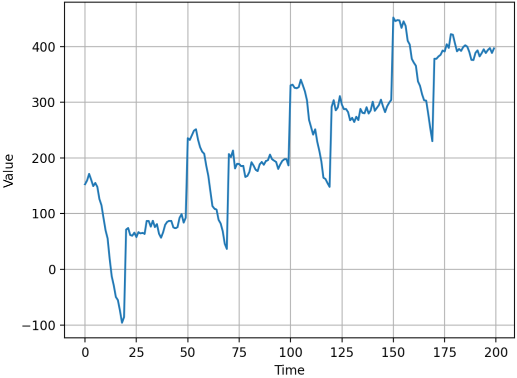
Auto-correlation

- Deterministic type of decay: it correlates with a delay copy of itself (lag)



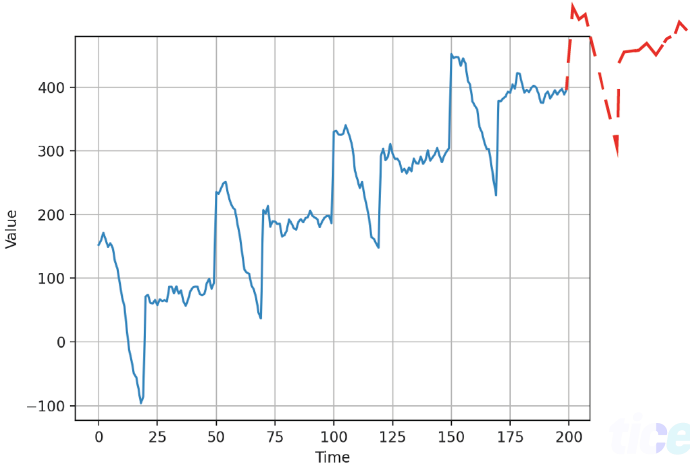
Trend + Seasonality + Auto-Correlation + Noise

- Trend + Seasonality + Auto-correlation + Noise



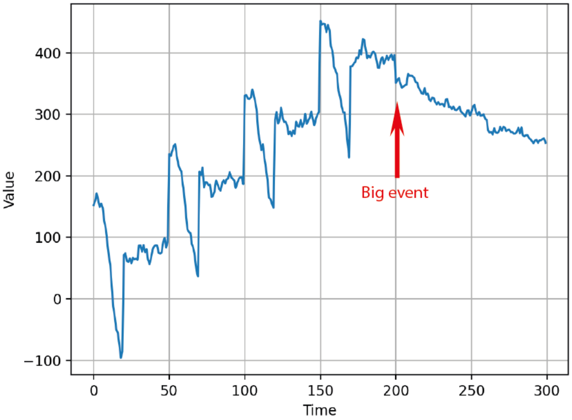
Forecast learned patterns

- A ML model is design to spot patterns. And then, predictions.



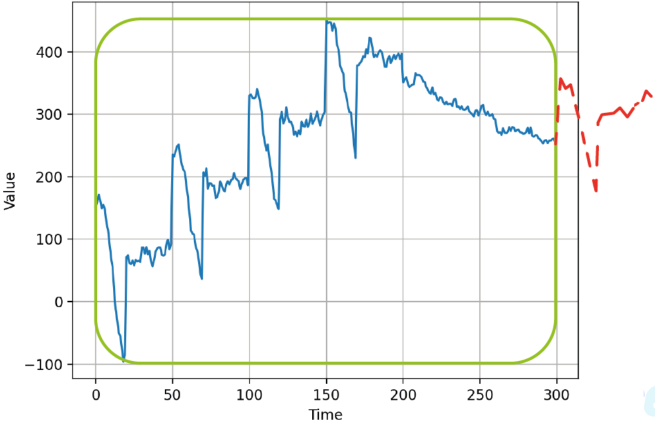
Non-stationary time series

- Time series behaviour can change drastically over time



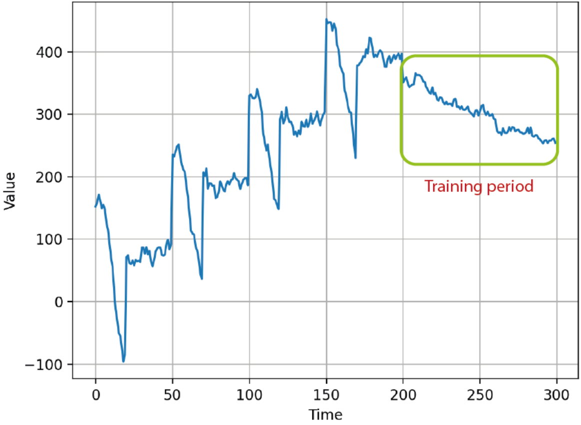
Non-stationary time series

- Difficult to perform



Non-stationary time series

- To predict: Train over a limited period of time.



Trading Strategies

- Long-term investment strategies
- **Short-term investment strategies**

Trading Strategies

- **High-frequency trading (HFT)**
  - Very short-term investment strategies
  - High volumes and high speeds.
  - Hold a position for **minutes** or **seconds**.
  - **60–73%** of all US equity trading volume.

1 Introduction

2 Materials

3 Methodology

4 Experiments

5 Results

6 Conclusions

Materials

- **Programming Language:** Python
- **Library:** Tensorflow
- **Hardware:** Google Colaboratory (Colab)
  - GPU: Nvidia K80s
  - RAM: 12.69GB
  - Disk: 68.35GB

1 Introduction

2 Materials

3 Methodology

Build Dataset

Model Setup

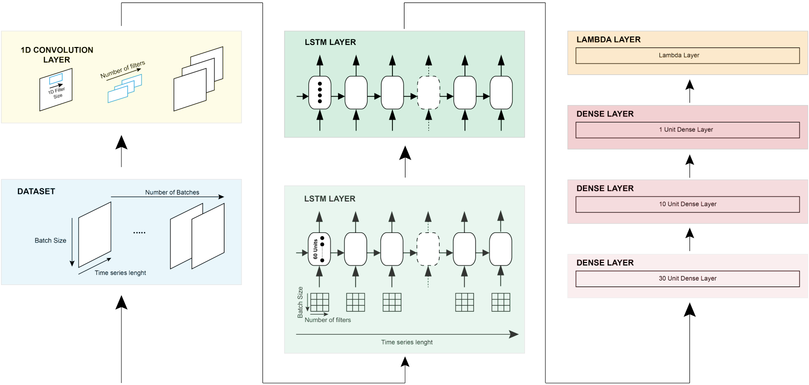
Training Parameters

4 Experiments

5 Results

6 Conclusions

Model Architecture

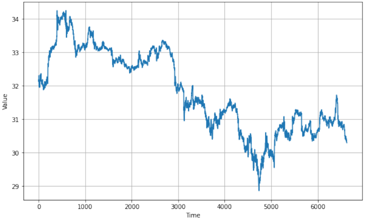


Data description

- **Intraday** stock prices of Amazon (ticker: AMZN)
- Last **60** days (January 25 - March 25 of 2021)
- **2m & 5m** Intervals
- Volume, dividends, **open**, high, low, and close price.
- Data source: **yfinance** (Yahoo! finance)

Univariate time series

- **Opening prices**
- **2m Intervals:** 6000-6500 observations
- **5m Intervals:** 3000-3500 observations



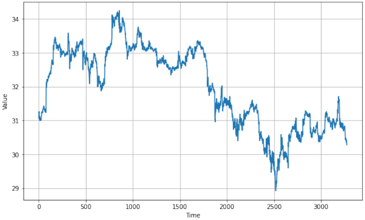


Fig 1: **Left**: 2 minute intervals - **Right**: 5 minute intervals

Split data for training and test

- **2m:** 3500 observations
- **5m:** 2000 observations

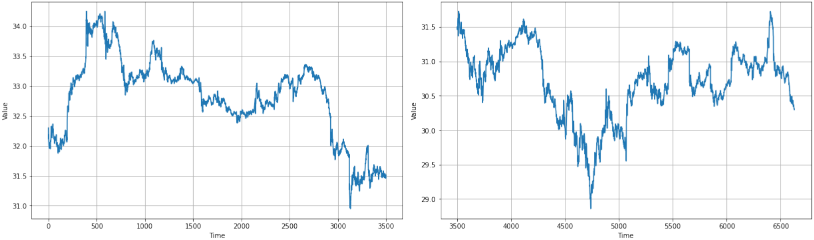
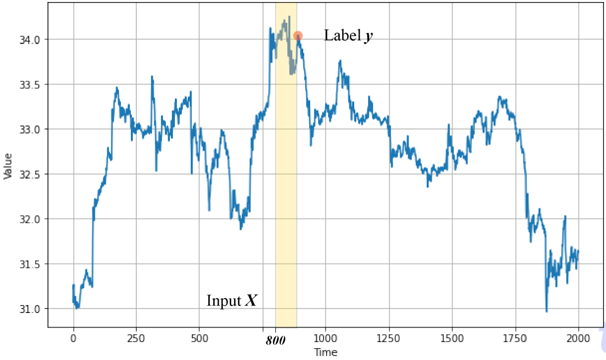


Fig 2: **2m:** 0-3500 (Training) 3501-6000 (Test)

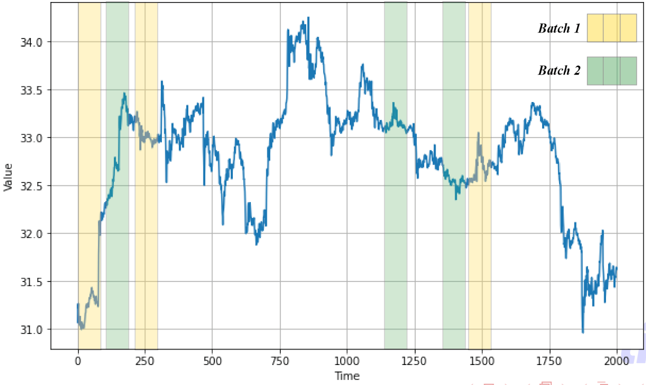
Define inputs and labels for the training dataset

- **2m Intervals**
  - Window size: 60 data points (2h)
- **5m Intervals**
  - Window size: 108 data points (9h)



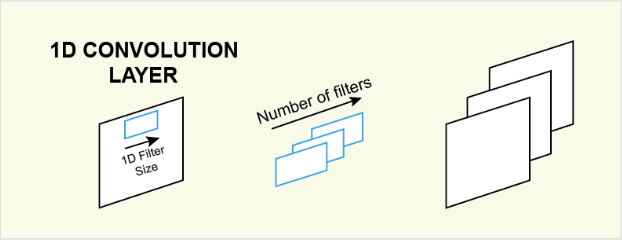
Batching the dataset

- **Shuffling the data:** Reduce the variance and the possibility of overfitting
- **Batch data:** Merge randomly selected sequences into a batch for training



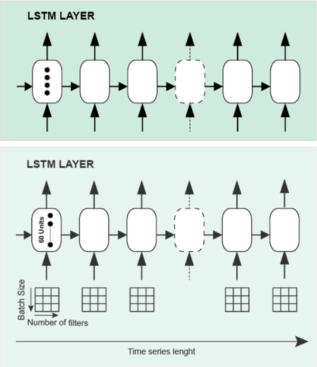
Convolutional Neural Network

- CNN setup:
  - **Layers:** 1
  - **Filters:** 60
  - **Kernel Size:** 5
  - **Strides:** 1
  - **Padding:** Causal (1D convolutions)



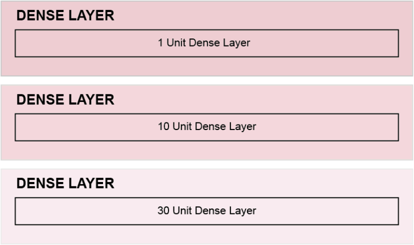
Long short-term memory (LSTM)

- **Layers:** 2
- **Hidden units:** 60
- **Returning Sequence:** True
- **Activation function:** tanh
- **Recurrent activation:** sigmoid



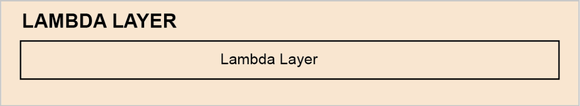
Dense Layers

- Third Dense Layer **Output layer**:
  - **Hidden units:** 1
- Second Dense Layer:
  - **Hidden units:** 10
  - **Activation Function:** ReLU
- First Dense Layer:
  - **Hidden units:** 30
  - **Activation Function:** ReLU



Lambda Layer

- Lambda Layer:
  - **Lambda function:** *lambda* $x : x * 400$
  - Applied to scale up the output values of the last dense layer in order to help the learning.

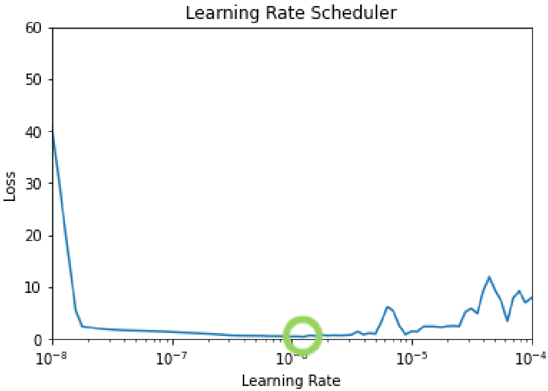


Fine-tuning parameters

- **Loss Function:** Huber
  - Less sensitive to outliers
- **Optimizer:** Stochastic Gradient Descent (SGD)
  - Iterative method for optimizing an objective function with suitable smoothness properties

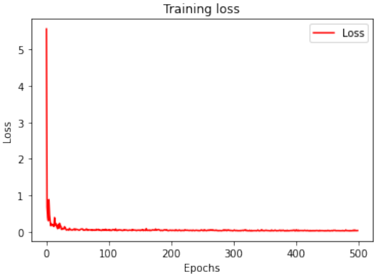
Fine-tuning parameters

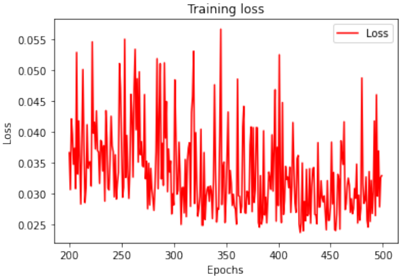
- **Learning Rate (LR)**
  - Define the optimum learning rate for SGD
  - Modulate how the LR of the SGD changes over time
  - Loss vs. Learning Rate



Fine-tuning parameters

- **Training Epochs**
  - Compare Loss vs. Training Epochs

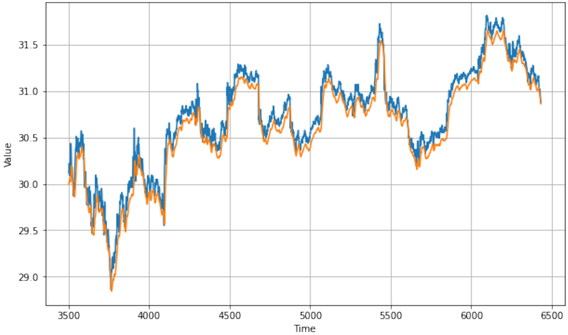




Varying window sizes 2m

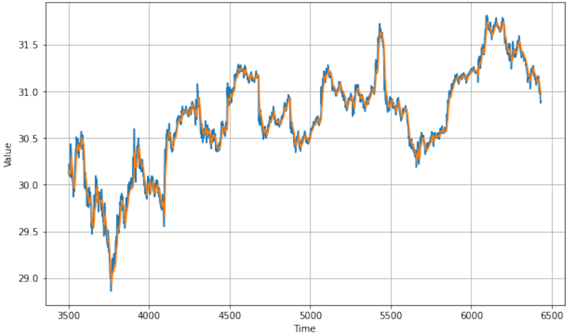
- Define an **optimum window size** in order to face the characteristic of non-stationary time series

- Window size: 40 - Intervals: 2m
- Forecasted (Orange) vs. Actual values (Blue)



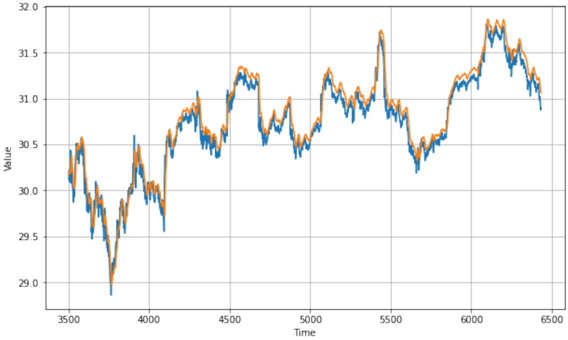
Forecasting Approach	Interval	WS	MAE
One-step	2 m	40	11.53

- Window size: 60 - Intervals: 2m
- Forecasted (Orange) vs. Actual values (Blue)



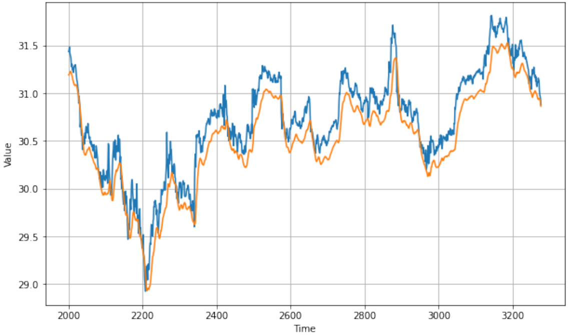
Forecasting Approach	Interval	WS	MAE
One-step	2 m	60	5.7

- Window size: 80 - Intervals: 2m
- Forecasted (Orange) vs. Actual values (Blue)



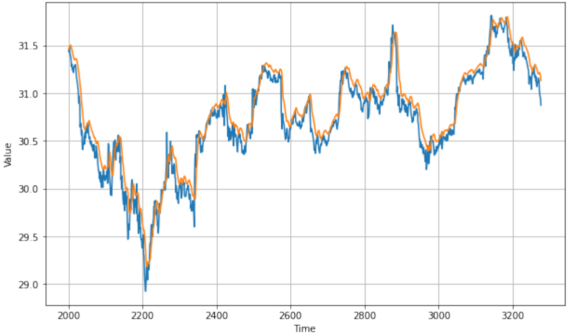
Forecasting Approach	Interval	WS	MAE
One-step	2 m	80	9.49

- Window size: 46 - Intervals 5m
- Forecasted (Orange) vs. Actual values (Blue)



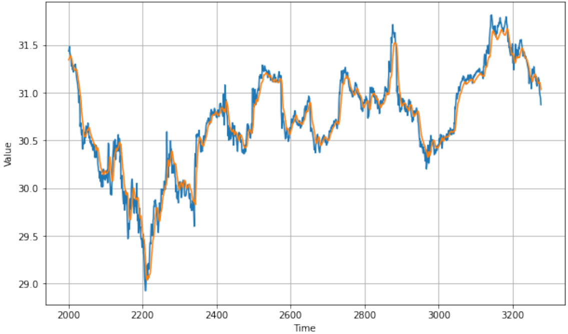
Forecasting Approach	Interval	WS	MAE
One-step	5 m	46	11.57

- Window size: 72 - Intervals 5m
- Forecasted (Orange) vs. Actual values (Blue)



Forecasting Approach	Interval	WS	MAE
One-step	5 m	72	10.32

- Window size: 108 - Intervals 5m
- Forecasted (Orange) vs. Actual values (Blue)



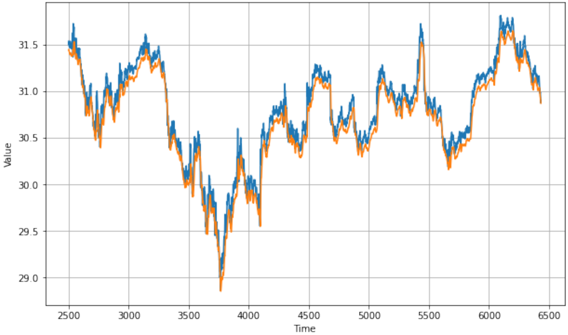
Forecasting Approach	Interval	WS	MAE
One-step	5 m	108	9.11

1. Introduction
2. Materials
3. Methodology
4. Experiments
   - Varying window sizes
   - Varying number of training examples
5. Results
6. Conclusions

Varying number of training examples 2m

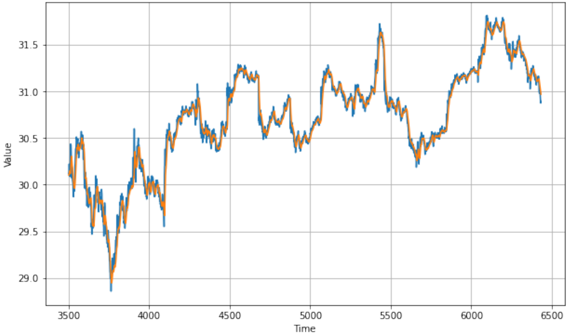
- Performance of the model is affected by varying the number of training examples
- More training examples does not always increase the accuracy of the model

- Training examples: 2500 - Interval: 2m
- Forecasted (Orange) vs. Actual values (Blue)



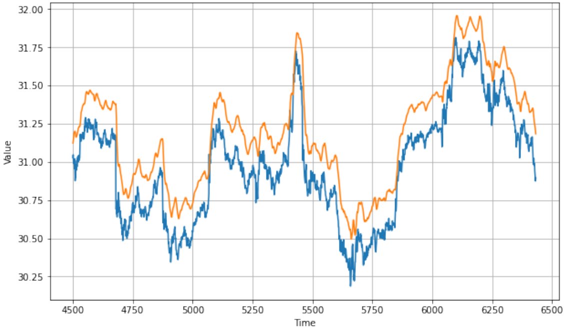
Forecasting Approach	Interval	TE	MAE
One-step	2 m	2500	8.83

- Training examples: 3500 - Interval: 2m
- Forecasted (Orange) vs. Actual values (Blue)



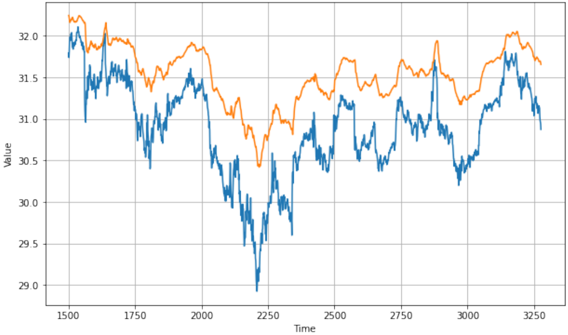
Forecasting Approach	Interval	TE	MAE
One-step	2 m	3500	5.7

- Training examples: 4500 - Interval: 2m
- Forecasted (Orange) vs. Actual values (Blue)



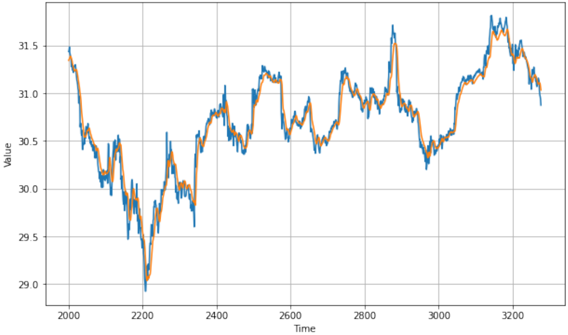
Forecasting Approach	Interval	TE	MAE
One-step	2 m	4500	21.99

- Training examples: 1500 - Interval: 5m
- Forecasted (Orange) vs. Actual values (Blue)



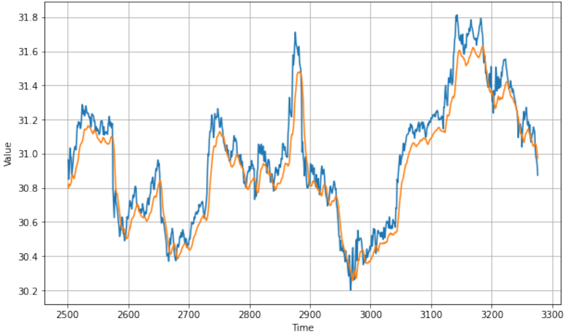
Forecasting Approach	Interval	TE	MAE
One-step	5 m	1500	6.46

- Training examples: 2000 - Interval: 5m
- Forecasted (Orange) vs. Actual values (Blue)



Forecasting Approach	Interval	TE	MAE
One-step	5 m	2000	9.11

- Training examples: 2500 - Interval: 5m
- Forecasted (Orange) vs. Actual values (Blue)

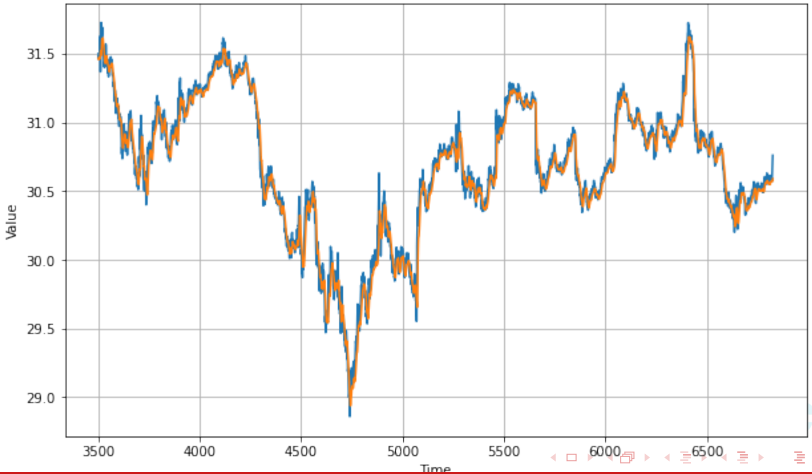


Forecasting Approach	Interval	TE	MAE
One-step	5 m	2500	9.52

1. Introduction
2. Materials
3. Methodology
4. Experiments
5. Results
   - Forecasting
   - MAE for one-step and multi-step forecasting
6. Conclusions

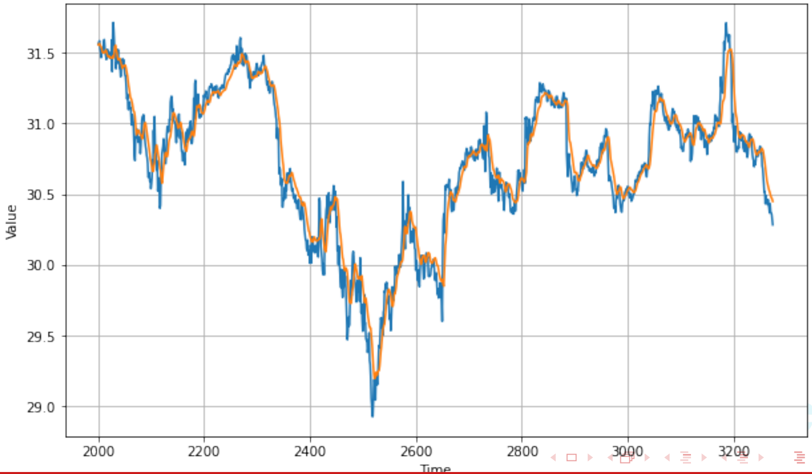
One-step forecasting

- **2-minute intervals**
- Forecasted (Orange) vs. Actual values (Blue)



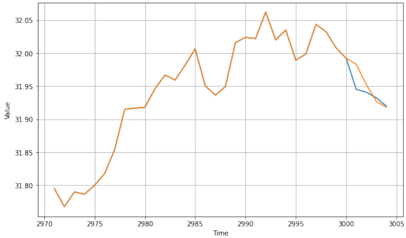
One-step forecasting

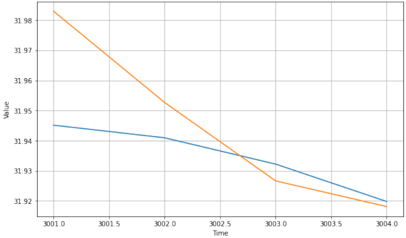
- **5-minute intervals**
- Forecasted (Orange) vs. Actual values (Blue)



Four-step forecasting

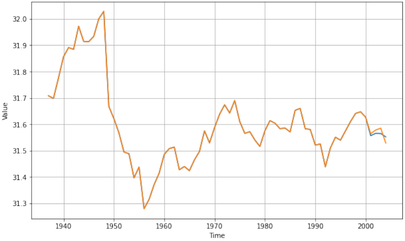
- **2-minute intervals**
- Forecasted (Blue) vs. Actual values (Orange)

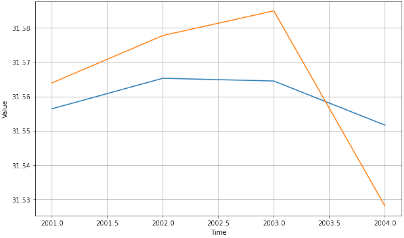




Four-step forecasting

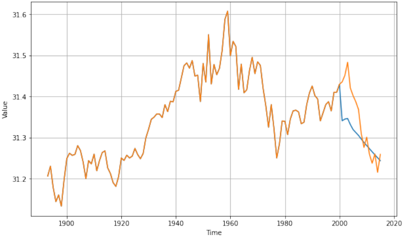
- **5-minute intervals**
- Forecasted (Blue) vs. Actual values (Orange)

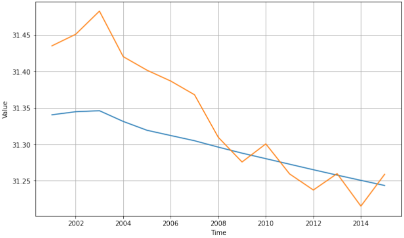




Fifteen-step forecasting

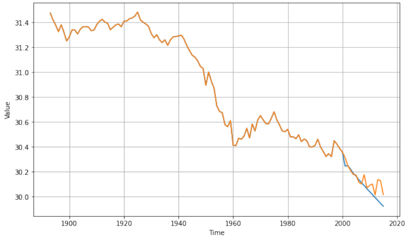
- **5-minute intervals**
- Forecasted (Blue) vs. Actual values (Orange)
- March 24, 2021.

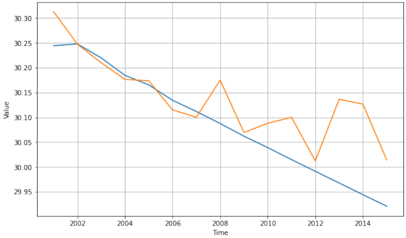




Fifteen-step forecasting

- **5-minute intervals**
- Forecasted (Blue) vs. Actual values (Orange)
- March 25, 2021.





1. Introduction
2. Materials
3. Methodology
4. Experiments
5. Results
    - Forecasting
    - MAE for one-step and multi-step forecasting
6. Conclusions

One-step and multi-step forecasting

Forecasting Approach	Interval	WS	TE	MAE
One-step	2 m	60	3500	6.7
One-step	5 m	108	2000	9.94
Four-step	2 m	60	3500	3.49
Four-step	5 m	108	2000	8.07
Fifteen-step	5 m	108	2000	9.84

- Combination of different deep architectures improves the capability of the model for identifying interrelations within the time series
- Deep architectures can be applied successfully
  - Accuracy, and forecasting speed
- More data does not increase the accuracy of the model
  - Accurate window size
  - Accurate number of training examples
- Future Work
  - Multivariate dataset
  - Apply high-performance computing (HPC) techniques for training the model
  - Develop trading strategies based on fundamental and technical analysis

Thanks!