

Consumer price index prediction using Long Short Term Memory (LSTM) based cloud computing

S Zahara¹, Sugianto¹ and M B Ilmiddaviq¹

¹Universitas Islam Majapahit, Jl. Raya Jabon Km 07 Mojokerto, Jawa Timur, Indonesia

Email: soffazahara@unim.ac.id

Abstract. Long Short Term Memory (LSTM) is known as optimized Recurrent Neural Network (RNN) architectures that overcome the lack of RNN's about maintaining long period of memories information. As part of machine learning networks, LSTM also notable as the right choice for time-series prediction. Inflation rate has been used for decision making for central banks also private sector. In Indonesia, CPI (Consumer Price Index) is one of best practice inflation indicators besides Wholesale Price Index and The Gross Domestic Product (GDP). Since CPI data could be used as a direction for next inflation move, we conducted CPI prediction model using Long Short Term Memory Method. The network model input consists of 34 variables of staple price in Surabaya and the output is CPI value. In the interest of predictive accuracy improvement, we used several optimization algorithm i.e. Stochastic Gradient Descent (sgd), Root Mean Square Propagation (RMSProp), Adaptive Gradient (AdaGrad), Adaptive moment (Adam), Adadelta, Nesterov Adam (Nadam) and Adamax. The result indicate that Nesterov Adam has 4.088 RMSE's value, less than other algorithm which indicate the most accurate optimization algorithm to predict CPI value.

1. Introduction

The use of deep learning for the purposes of building prediction models has been widely used by experts to improve the accuracy of predictions in various sectors such as economics, health, sports, education, agriculture and animal husbandry. Currently, it plays role in economic data analysis since the increasingly numbers of data recorded about consumers and purchasing behavior. Prediction in the economic sector turn into crucial state when accurate prediction results can support the government in determining the next policy. Economics indicator that serve information about the prices of goods and services paid by consumers known as Consumer Price Index (CPI). CPI is one of economic parameters issued by the Statistics Indonesia (BPS) intend to inform purchase prices at the consumer level [1]. The movement of the purchase price at the consumer level from time to time is the basis for calculating the CPI.

In Indonesia, the CPI value issued every month and year is the most frequently used inflation indicator in addition to the Large Trade Price Index (IHPB), Producer Price Index (IHP), Gross Domestic Product Deflator (GDP), and Asset Price Index [1]. CPI is chosen as one of the inflation references due to the frequency and timeliness produced. In addition to inflation which is influential in the government's decision to take economic policy, it is also used as a consideration for various aspects of social finance such as retirement, unemployment, and government funding [2].

Deep learning network is known have good system performance and produce an optimal system to processing data in machine learning and big data world. Deep learning is a sub-section of machine learning that uses many layers as a non-linear information processing media [3]. One type of deep learning algorithm is Long Short Term Memory (LSTM) which is a development of the Recurrent Neural Network (RNN) algorithm which overcomes the main problem of RNN, namely not being able to process sequential information in the long run, especially time-series data processing [4].



Gao [5] built modeling for stock market prediction used 4 different methods, namely moving average (MA), exponential moving average (EMA), Support Vector Machine (SVM) and LSTM. From the test results proved that LSTM had the highest accuracy value compared to other methods. In addition, Jeenanunta as in [6] had also proven in his research that compared stock predictions using the LSTM method with 3 types of predictions of Thai shares, namely CPALL, SCB, and KTB. LSTM managed to outperform with an error performance value of less than 2%. There were certain studies on CPI prediction in Indonesia, namely Dewi [7] used SVR and MAPE as a measure of accuracy to predicted CPI using the housing group for water, electricity, gas, and fuel. The research above only used CPI variables as both input and output prediction models made. Budiastuti [8] also predicted the Daily Consumer Price Index using the Support Vector Regression (SVR) method with 4 types of kernel variations.

Food price data is sequential and historical, as well as CPI data become one of the parameters of inflation that can influence policy any time so that this study aims to predicted the monthly CPI index using the LSTM algorithm. Model input data used price data of 34 types of staple food in East Java originating from Department of Industry and Commerce of East Java Province. While for predictive output used monthly CPI data obtained from Statistics Indonesia. In order to improved accuracy for making prediction models, this study used 7 different optimization algorithms, namely Stochastic Gradient Descent (SGD), Root Mean Square Propagation (RMSProp), Adaptive Gradient (AdaGrad), Adaptive moment (Adam), Adadelata, Nesterov Adam (Nadam) and Adamax.

The develoment of prediction model in machine learning particularly deep learning acclaimed as time yet memory consuming [9][10], meanwhile appearence of Cloud Computing environment offered best alternative to accelerate computation time. Nowadays, abundants of Cloud Computing providers offered computing environment that suited machine learning requirements very well such as Amazon Web Services(AWS), Google Cloud, Microsoft Azure. One of the advantages that can be obtained from the use of Cloud Computing for is that users do not need to think about the complexity of the infrastructure if the development of the system becomes more complex and large. Amazon Web Service in one of largest Cloud Computing providers that have dozens of great services too fasilitate all user needs, starts from VPS(Virtual Private Server), Cloud Storage, Business Intelegant Platform, IoT(Internet of Things), also Machine Learning Platform and so on. AWS also provides complete documentations also support forum to help customers. Therefore, AWS became leader in public Cloud market since the comprehensiveness of their service and the product realibility. In the terms of Deep Learning platform, AWS supports AI(Artificial Intelegant) model experiment using TensorFlow, PyTorch, Apache MXNet, Chainer, Gluon, Horovod, and Keras. The construction of prediction models in this study utilized the Amazon Web Service (AWS) EC2 virtual computer to facilitate the retrieval of raw data as well as reduce computing time. EC2 (Elastic Compute Cloud) is one of the AWS services that provides virtual computing that can be configured according to user needs [11].

2. Methodology

Figure 1 depicts the stage of this research divided into 5 stages, namely the raw data retrieval stage then the pre-processing process, that is the process of cleaning and transforming data, afterwards the data is allocated into training data and testing data. The next stage is the development of prediction models using LSTM method. After the prediction model is built, accuracy performance evaluation is carried out by compared 7 types of optimization algorithms. The method proposed in this paper was adapted from Thakur et al [12] method with improvement in Cloud Computing environment.

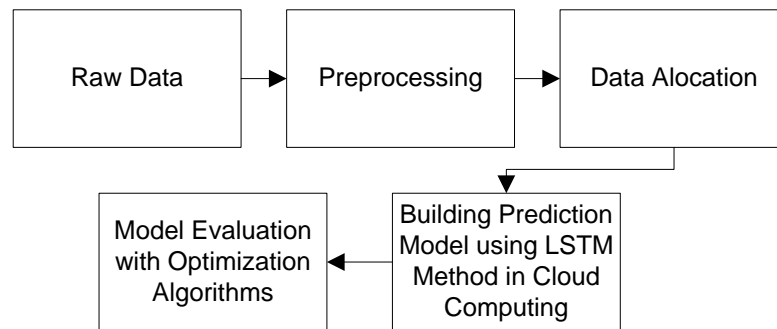


Figure 1. Research methodology.

2.1. Raw Data Retrieval

The data used in this study divided into 2 categories, namely input and output data. The input data came from 34 types of staple food prices in 38 cities in East Java. It were obtained online from the Department of Industry and Commerce of East Java Province website, with taking periods from 2014 to 2018. In this research only the price of staple foodstuffs was used because food price data frequently fluctuated making it more suitable to be used as data for predictions. Furthermore, the output data were CPI value that issued regularly every month by the Statistics Indonesia particularly in East Jawa Region for the same collection period from 2014 to 2018.

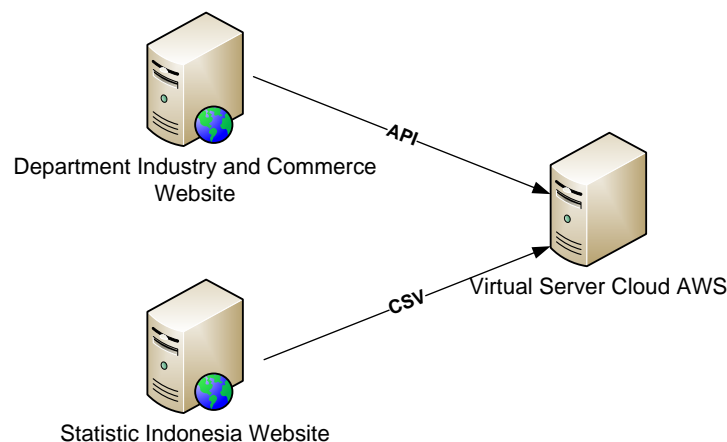


Figure 2. Prediction model architecture.

Figure 2 illustrates the data collection techniques used in the construction of prediction models. Staple food prices were obtained from Department of Industry and Commerce of East Java Province's website, www.siskaperbapo.com through the API (Application Programing Interface) that had been provided. Furthermore, for the CPI data obtained from BPS, it is obtained manually downloading the CSV format file.

2.2. Preprocessing

The preprocessing stage divided into 2 stages, missing data processing stage and the scaling stage. In the process of retrieved staple price data, there were several lost daily price data, the solution to the problem was to filled in the data gaps with the average price of the previous day.

Since missing data problems has been resolved, the next step is rescaling to min max normalization data. Scaling technique intend to tranform data value smaller in scale without changing the information inside. This technique is used to overcome the problem of a considerable value gap between the staple

data and CPI data. Scaling is done by converting to scale (0,1) [13] which is the lower limit of 0 and the upper limit of 1.

2.3. Data Allocation

Data allocation divides the dataset into 2 types, namely training data or learning data and test data or trial data. In the process of building prediction model, the distribution of training data and testing data is 70:30.

2.4. Prediction Model Building Using LSTM Method

The development of the prediction model uses the python programming language and several supporting libraries for deep learning, which are scikit-learned, keras, pandas, numpy, matplotlib and tensorflow as backend.

LSTM model architecture used consists of sequential input layers means a stack of linear layers consisting of 34 input variables. Figure 3 depicts the input data were processed into the LSTM network architecture with a configuration of 1 hidden layer consisting of 50 neurons. The number of epochs used is 50 and the batch size used is 72. Furthermore, the data were streamed to the output dense layer with 1 output variable. All layer transfers use fully-connected networks

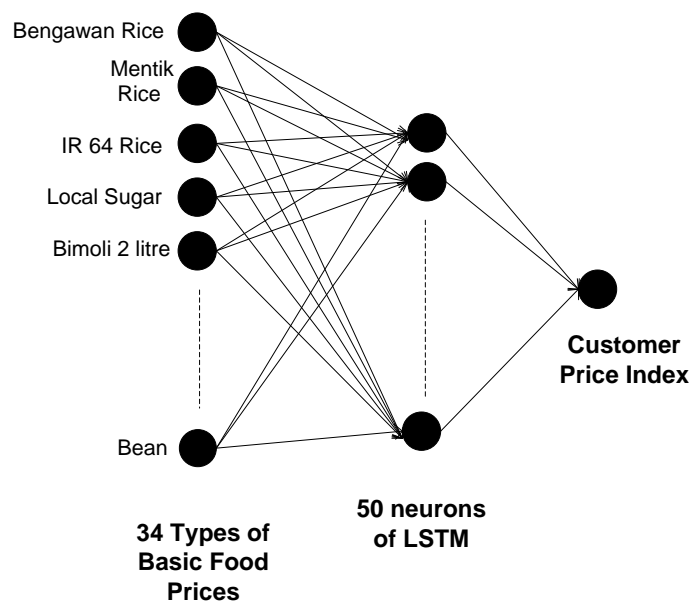


Figure 3. Prediction model architecture.

LSTM is one of the latest developments of the RNN (Recurrent Neural Network) algorithm. While an ordinary simple RNN architecture consisting of a recurrent network module containing a simple tan layer function structure, the LSTM network is described as several repetitive chain modules where in each module there is another sub module containing sigmoid gate functions. Gate function in LSTM mostly consist of 3 gates specifically input gate, forget gate, and output gate. The amount of gate could be vary depend on the architecture that has been used. The sigmoid function is a function that regulates the amount of information that is passed.

In the process of calculating LSTM, input data in the form of time series is assumed to be $x=(x_1, x_2, \dots, x_n)$ and output $y=(y_1, y_2, \dots, y_n)$ then the LSTM calculation process is done with the following formulas :

$$h_t = H(W_{hx}x_t + W_{hh}h_{t-1} + b_h) \quad (1)$$

$$p_t = W_{hy} y_{t-1} + b_y \quad (2)$$

The W symbol is the weight of the matrix and b is biased in the vectors form. Furthermore, hidden state calculation is done as follows:

$$i_t = \sigma(W_{ix} x_t + W_{hh} h_{t-1} + W_{ic} c_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma(W_{fx} x_t + W_{fh} h_{t-1} + W_{fc} c_{t-1} + b_f) \quad (4)$$

$$c_t = f_t * c_{t-1} + i_t * g(W_{cx} x_t + W_{ch} h_{t-1} + W_{cc} c_{t-1} + b_c) \quad (5)$$

$$o_t = \sigma(W_{ox} x_t + W_{oh} h_{t-1} + W_{oc} c_{t-1} + b_o) \quad (6)$$

$$h_t = o_t * h(c_t) \quad (7)$$

The * symbol means the product of multiplication between 2 matrices. Moreover, g and h are sigmoid functions with a range of values between [-2, 2] and [-1, 1]. σ is a standard sigmoid function also e is square loss function. The calculation of sigmoid and loss function is done with the following equations :

$$\sigma(x) = \frac{1}{1+e^x} \quad (8)$$

$$e = \sum_{t=1}^n (y_t - p_t)^2 \quad (9)$$

2.5. Model Evaluation with Optimization Algorithm

After the prediction model successfully built, the training process will be carried out until it reaches the optimum accuracy. This process uses certain kinds of optimization algorithms, namely Stochastic Gradient Descent (SGD), Root Mean Square Propagation (RMSProp), Adaptive Gradient (AdaGrad), Adaptive moment (Adam), Adadelta, Nesterov Adam (Nadam) and Adamax. SGD is one of three types original gradient descent, meanwhile the rest of other algorithms such as RMSProp, AdaGrad, Adam, Adadelta, Nadam, and Adamax are the latest development version of gradient descent[14].

Stochastic Gradient Descent(SGD) performance is always change parameters in each data that is trained. In updating parameters, SGD does not loop thus it is faster for large datasets. Meanwhile, Adagrad is a gradient-based optimization algorithm that adapts the learning rate variable into parameters. The more parameters updated, the smaller the learning rate will be. Adagrad inherits SGD reliability properties, and is suitable for large scale data[14].

Adadelta algorithm is the development version of Adagrad that reduce the aggressive nature of Adagrad in reducing the learning rate value. In addition, the performance and objectives of the RMSProp algorithm is resemble to Adadelta, yet RMSProp can outperform Adagrad in reducing the learning rate value. Adaptive Moment Estimation (Adam) is an optimization method that calculates the level of learning adaptively for each parameter. Like Adadelta and RMSProp, Adam keeps the previous process gradient average exponentially. Nadam (Nesterov-accelerated Adaptive Moment Estimation) is a combination of RMSProp and Nesterov accelerated gradient (NAG).

RMSE (Root Mean Square Error) is used as a parameter of accuracy and efficiency of each optimization algorithm. RMSE calculates the error value or the difference between predictive and actual data. N denotes total of data then the Y_i is predicted data also Y is actual data.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - Y)^2} \quad (10)$$

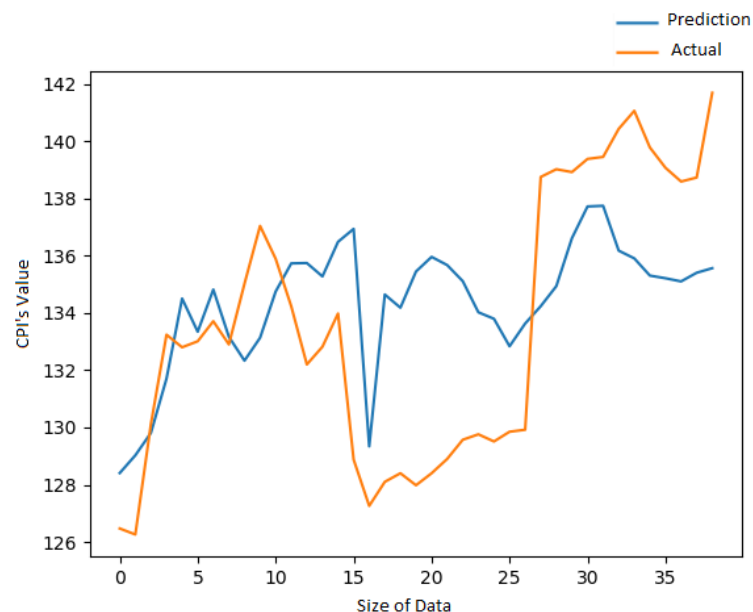
3. Results and discussion

Based on the the experiment that have been done the result below in Table 1 describe about comparison prediction accuracy between 7 optimization algorithms. Due to abundant of missing staple food price variables in several town, we decided to only used single town that is Surabaya for subject of this research. The LSTM model predict CPI at the moment (t) using previous data (t-1).

Table 1. Performance Result.

No	Optimization	RMSE value
1	Stochastic Gradient Descent (SGD)	4,219
2	Root Mean Square Propagation (RMSProp)	6,119
3	Adaptive Gradient(AdaGrad)	5,435
4	Adaptive moment(Adam)	6,926
5	Adadelata	5,207
6	Nesterov Adam(Nadam)	4,088
7	Adamax	5,835

Table 2 informs about comparison accuracy value between 7 different optimization algorithm applied in proposed model. The most accurate performance in this model belongs to Nesterov Adam that has the smallest RMSE value among others that is 4.088, followed by SGD, Adadelata, Adagrad, Adamax, RMSProp, and Adam in sequence. Among other optimization algorithms, Nesterov Adam(Nadam) is the newest algorithm combining RMSProp and Adam. The result proved that Nadam as newest optimization offer great performance compared previous algorithm to predicted CPI value.

**Figure 4.** Plot Graph Result Using Nadam

The plot graphs above on Figure 4 illustrate prediction and actual ratio in CPI prediction model in Nadam algorithm. The orange graph shows CPI actual data, while blue graph deals with CPI prediction value that obtained from testing data. Moreover, the x axis informs the size of testing data then the opposite y axis shows about CPI value. The size of testing data were range from 1 to 40 then range of CPI values were from 126 to 142. The According to the graph, there were some point that show predicted value successfully approaching forecast CPI actual data.

Table 2. Comparison Result

Period	Actual CPI	Predicted CPI
May 2016	132.9	128.0594
June 2016	135.02	129.2686
July 2016	137.04	130.3162
August 2016	135.89	132.2575
September 2016	134.23	135.2031
October 2016	132.2	133.1354
November 2016	132.82	132.8289
December 2016	133.98	133.1354
January 2017	128.88	132.8289
February 2017	127.26	133.9902
March 2017	128.1	135.3659
April 2017	128.4	135.6953
May 2017	127.98	135.8108
June 2017	128.4	135.5442
July 2017	128.9	136.0939
August 2017	129.57	137.0477
September 2017	129.76	129.4482
October 2017	129.51	135.3687
November 2017	129.85	134.9568
December 2017	129.92	136.0382
January 2018	138.75	136.2948
February 2018	139.02	136.2935
March 2018	138.92	135.9997
April 2018	139.38	135.9697
May 2018	139.45	136.0438
June 2018	140.43	134.831
July 2018	141.06	135.6731
August 2018	139.78	136.6882
September 2018	139.07	137.2478
October 2017	138.59	137.2249
November 2018	138.73	137.3233
December 2018	141	137.4724

Table 2 informs the comparison of predicted and actual CPI values using the results of Nadam optimization algorithm. Predicted CPI value in December 2016 was 133.98, which has closest to the actual CPI value that is 133.98. From the all result that have been obtained, RMSE value that portrayed evaluation model in this research still relatively large. It can be seen LSTM method still not yet called the best method to predict CPI value. There were many factors that influence among others, significant CPI and food staple price value that changes especially in certain months, for example during holidays and new years. In addition, the calculation of the base year of the CPI that changed following the base year of Survey Biaya Hidup (SBH) was also very influential.

4. Conclusion

The application of CPI value for inflation measurement was widely used by most countries since it could portrayed the real economics situation for example staple food price movement. Therefore, an

accurate prediction is needed to deal with undesired things as well as immediately anticipated by government.

In this research we conducted monthly CPI prediction model based LSTM method of deep learning with non linier parameter input that were daily staple food prices in Surabaya. We applied several optimization algorithms to obtained best accuracy in LSTM method applied. The result showed that Nadam (Nesterov Adam) as the latest algorithm proved the best performance among others as the RMSE value is 4.088. Even the accuracy in this model was still far from expectations, there were several aspect of evaluation could be implemented in the next research. The total variation of epoch, hidden layer, batch size, input variable could be tested in future work to obtained optimum accuracy. This research could be used as reference for predicting inflation rate using Costumer Price Index value while the hint of CPI data were from staple food price.

5. References

- [1] D. Statistik, 2016, *Metadata Indeks Harga Konsumen*, [Online] (Updated March 2016) Available at : [https://www.bi.go.id/id/statistik/metadata/seki/Documents/12.Inflasi-Indeks Harga Konsumen \(IHK-IND\)2016.pdf](https://www.bi.go.id/id/statistik/metadata/seki/Documents/12.Inflasi-Indeks%20Harga%20Konsumen%20(IHK-IND)2016.pdf). [Accesses 10 June 2019]
- [2] ILO/IMF/OECD/UNECE/Eurostat/The World Bank. 2004. *Consumer price index manual: Theory and Practice*. 1st ed. Geneva : Publications Bureau
- [3] L. Deng, O. M. Way, D. Yu, and O. M. Way. 2014. Deep Learning : Methods and Applications. *Foundations and Trends in Signal Processing*. vol. 7. no. 2013. pp. 197–387.
- [4] Hochreiter and J. Schmidhuber. 1997. Long short-term memory. *Neural Computation*. vol. 9. no. 8. pp. 1735–1780.
- [5] T. Gao, Y. Chai, and Y. Liu. 2018. Applying Long Short Term Momory Neural Networks for Predicting Stock Closing Price. *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*. pp. 3–6. Beijing, China 24-26 Nov. 2017
- [6] C. Jeenanunta, R. Chaysiri, and L. Thong. Stock Price Prediction With Long Short - Term Memory Recurrent Neural Network. *2018 Int. Conf. Embed. Syst. Intell. Technol. Int. Conf. Inf. Commun. Technol. Embed. Syst.*, pp. 1–7. Khon Kaen, Thailand 7-9 May 2018
- [7] K. Dewi, P. P. Adikara, and S. Adinugroho. *Prediksi Indeks Harga Konsumen (IHK) Kelompok Perumahan , Air , Listrik , Gas Dan Bahan Bakar Menggunakan Metode Support Vector Regression*. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* vol. 2, no. 10, pp. 3856–3862, 2018.
- [8] I. A. Budiastuti, S. M. S. Nugroho, and M. Hariadi. Predicting daily consumer price index using support vector regression method. *QiR 2017 - 2017 15th Int. Conf. Qual. Res. Int. Symp. Electr. Comput. Eng.* Nusa Dua, Indonesia 24-27 July 2017
- [9] L. Feng, P. Kudva, D. Da Silva, and J. Hu, “Exploring Serverless Computing for Neural Network Training,” pp. 334–341, 2018.
- [10] S. Solutions, “Machine Learning and Cloud Computing: Survey of Distributed and SaaS Solutions *,” 2012.
- [11] A. W. Services. *Amazon Elastic Compute Cloud User Guide for Linux Instances*. [Online] (Updated 2019) Availabale at : <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-ug.pdf> [Accesses 20 Juni 2019]
- [12] R. Fu, Z. Zhang, and L. Li. Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction. *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. Wuhan, China 11-13 Nov. 2016
- [12] G.S.M. Thakur, R. Bhattacharyya, and S. S. Mondal, “Artificial Neural Network Based Model For Forecasting of Inflation in India”. *Fuxxy Information and Engineering*. pp. 87–100, 2016..
- [13] S. G. K. Patro and K. Kumar, “Normalization : A Preprocessing Stage. [Online](Updated 19 March 2015) Available at : <https://arxiv.org/abs/1503.06462>. [Accesses 19 Juni 2019]
- [14] S. Ruder. An overview of gradient descent optimization [Online](Updated 19 Jan 2016) Available at : <https://arxiv.org/abs/1609.047477>. [Accesses 19 Juni 2019]