

Order number:
Serial number:

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research



**UNIVERSITY OF ECHAHID HAMA
LAKHDER - ELOUED**
FACULTY OF EXACT SCIENCES
Computer Science department



End of study memory

Presented for the Diploma of

ACADEMIC MASTER

Domain : **Mathematics and Computer Science**

Industry: **Computer Science**

Speciality: **Distributed Systems and Artificial Intelligence**

Presented by: FERHAT Adel

TOUMI Salsabil

Theme

**Exploiting LSTM Algorithm
For Renewable Energy**

Discussed on 06-22-2019 before the jury:

Mr: BOUCHERIT Ammar	MCA	President
Mr: ABBAS Messaoud	MAA	Reporter
Mr: NAOUI Med Anouar	MAA	Supervisor

Academic year 2018/2019

Thanks

Before all, we want to thank from the deep of our hearts, "ALLAH" who has all-powerful to give us strength, willingness and courage to do this work.

We would like to thank warmly and respectfully all those who contributed from near and far to the realization of this end of study project especially fathers and our families.

We want to thank our master framing Mr. NAOU Mohamed Anouar, Professor at the University of Echahid Hamma Lakhdar El-Oued, for proposing and supervising this topic for six months. We express our deep gratitude for making us benefit from his knowledge, but also from his methods of work, and especially his scientific rigor.

We thank the members of the jury for their generosity for acceptance of our work.

Our gratitude also goes to all those who collaborated in our study especially the teachers of the Department of Computer Science, University of Echahid Hamma Lakhdar El-Oued.

Also, to our colleagues from the 2018-2019 promotion. We also thank everyone who has participated from near or far in order to develop this work.

Abstract

Increased energy consumption in recent years based on oil extraction is about to end. Therefore, in contrast the growing interest in renewable energies, (never ending), has attracted a lot of research in this area in order to provide safe and adequate energy to the consumer.

As renewable energy production is still insufficient for consumption, on the other hand, it is closely related to weather or nature, so it is necessary to know how much we can predict future renewable energy production, which is challenging as it is very important for daily human consumption.

In this work, we used renewable energy sources, solar panels and wind tribunes with the Internet of things (IoT) technology to transfer the data and characteristics of energy generated and then deep learning. In order to propose an architecture to energy forecasting composed of three main modules, data extraction module, data storage module, machine Learning Module, in which we suggest a Long Short-Term Memory (LSTM) algorithm and compared it with three models to illustrate the power of our system that has the ability and preference to predict the size of energy production with the optimization of execution time .

Key words: Renewable Energy forecasting, Internet of Things, Deep learning, Long Short-Term Memory.

ملخص

استهلاك الطاقة المتزايد في السنوات الأخيرة المبني على استخراج البترول، بات على وشك النهاية، وعليه فإنه في المقابل تزايد الاهتمام بالطاقات المتجددة، التي جذبت الكثير من الأبحاث في هذا المجال من أجل توفير طاقة آمنة وكافية للمستهلك.

وعلى اعتبار أن إنتاج الطاقة المتجددة لا يزال غير كافي بالنسبة للاستهلاك من جهة، ومن جهة أخرى، ارتباطه الوثيق بحالة الطقس أو الطبيعة، لذلك بات من الضروري معرفة ما مدى قدرتنا على التنبؤ بإنتاج الطاقة المتجددة مستقبلا، وهو الأمر الذي يشكل تحديا باعتباره غاية في الأهمية بالنسبة للاستهلاك البشري اليومي.

في هذا العمل، استخدمنا مصادر الطاقة المتجددة، والألواح الشمسية، ومنابر الرياح مع تكنولوجيا إنترنت الأشياء لنقل بيانات وخصائص الطاقة المولدة ومن ثم التعلم العميق. من أجل اقتراح بنية للتنبؤ بالطاقة تتكون من ثلاث وحدات رئيسية، وحدة استخراج البيانات، وحدة تخزين البيانات، وحدة تعلم الآلة، نقترح فيها خوارزمية ذاكرة طويلة المدى ثم مقارنتها بثلاثة نماذج لتوضيح قوة نظامنا الذي لديه القدرة والأفضلية للتنبؤ بحجم إنتاج الطاقة مع تحسين وقت التنفيذ.

الكلمات المفتاحية: التنبؤ بالطاقة المتجددة، إنترنت الأشياء، التعلم العميق، الذاكرة طويلة-قصيرة الأجل.

Résumé

L'augmentation de la consommation d'énergie liée à l'extraction du pétrole est sur le point de s'achever. Par conséquent, l'intérêt croissant pour les énergies renouvelable sa attiré de nombreuses recherches dans ce domaine afin de fournir une énergie adéquate et sécurisé pour le consommateur.

La production d'énergie renouvelable reste insuffisante pour la consommation d'une part et son association étroite avec le climat ou la nature d'autre part, pour cette raison, il est devenu nécessaire de savoir à quel point nous pouvons prédire la production future d'énergie renouvelable, ce qui constitue un défi, Parce que c'est important Dans la consommation humaine quotidienne.

Dans ce travail, nous avons utilisé des sources d'énergie renouvelables, des panneaux solaires et tribunes du vent avec la technologie de l'Internet des objets pour transférer les données et les caractéristiques de l'énergie générée sous forme de données volumineuses et puis l'apprentissage en profondeur. Afin de proposer une architecture de prévision énergétique composée de trois modules principaux, module d'extraction de données, module de stockage de données, module d'apprentissage machine, dans le quelle nous proposons un algorithme de mémoire à court-long terme et l'a comparé à trois modèles pour illustrer la puissance de notre système qui a la capacité et la préférence de prédire la taille de la production d'énergie avec l'optimisation du temps d'exécution.

Mots clés: Prévision des énergies renouvelables, Internet des objets, apprentissage en profondeur, Mémoire à court-long terme.

Contents

Thanks.....	I
Abstract.....	II
ملخص.....	II
Résumé.....	III
General Introduction	1
Chapter 01: State of the art: renewable energy.....	2
State of the art: renewable energy	2
1 Introduction	3
2 Definition	3
3 Renewable energy types	3
3.1 Solar energy	3
3.1.1 Solar energy component.....	4
3.1.2 Advantages.....	5
3.2 Wind energy	5
3.2.1 Wind energy component.....	6
3.2.2 Advantages.....	7
4 Renewable energy comparison	7
5 Why use renewable energy	8
6 Conclusion	9
Chapter 02: Deep learning and big data system.....	10
Deep learning	10
1 Introduction	11
2 Deep learning	11
2.1 Definition	11
2.2 Architecture.....	11
3 Deep learning Algorithms	12

3.1	Back propagation	12
3.2	Dropout.....	12
3.3	Long Short-Term Memory (LSTM).....	13
3.3.1	LSTM characteristics.....	13
3.3.2	LSTM composition.....	14
4	Internet of Things IoT.....	15
4.1	Area of application.....	16
4.2	Platforms.....	16
4.2.1	Wearable platform.....	16
4.2.2	Embedded platform.....	16
5	Big Data.....	17
6	Conclusion.....	18
	Chapter 03: proposition of deep learning for renewable energy.....	19
	proposition of deep learning for renewable energy.....	19
1	Introduction.....	20
2	Related works.....	20
3	Proposition.....	21
3.1	Modular Architecture.....	21
3.2	Data Mining Process For Renewable Energy.....	23
3.3	Long Short Term Memory Algorithm.....	24
4	Validation and experimentation.....	26
4.1	Validation system.....	26
4.1.1	Development of LSTM algorithm.....	27
4.1.2	Description of the system functionality according to proposed architecture.....	27
4.2	Experimentation and result.....	28
4.2.1	results.....	28
4.2.2	Comparison with machine learning models.....	32
5	Conclusion.....	35

Chapter 04: implementation and development.....	36
implementation and development.....	36
1 Introduction.....	37
2 Used tools.....	37
2.1 Language of Python.....	37
2.2 Jupyter notebook.....	38
2.3 Libraries used.....	38
2.4 Hadoop big Data tool's.....	40
3 3 Programming Environment.....	41
3.1 Loading of Data set.....	41
3.2 Datasets Show.....	42
3.3 Elaboration of Data set.....	44
3.4 Train and Test Data Set.....	45
3.5 LSTM : create, fit and predictions.....	46
3.6 LR, KNN and CART algorithms.....	49
4 Conclusion.....	51
General conclusion.....	52
Bibliography.....	54

List of figures

FIGURE 1.1. CONCENTRATED SOLAR POWER	4
FIGURE 1.2. PHOTOVOLTAICS.....	4
FIGURE 1.3. SOLAR SYSTEM.....	5
FIGURE 1.4. WIND ENERGY.....	6
FIGURE 1.5. HORIZONTAL-AXIS WIND TURBINE	7
FIGURE 2.1. DEEP LEARNING & MACHINE LEARNING	11
FIGURE 2.2. DROP OUT.....	13
FIGURE 2.3. LSTM STRUCTURE.....	14
FIGURE 2.4. AREA OF INTERNET OF THINGS.....	15
FIGURE 2.5. THREE V FOR BIG DATA.....	17
FIGURE 3.1. MODULES OF THE SYSTEM.....	21
FIGURE 3.2. ARCHITECTURE OF OUR SYSTEM.....	22
FIGURE 3.3. DATA MINING PROCESS CRISP-DM	24
FIGURE 3.4. LSTM CELLS CONNECTED TO EACH OTHER	25
FIGURE 3.5. HADOOP FOR AUSTRALIA RENEWABLE ENERGY DATA	27
FIGURE 3.6. WIND POWER GENERATION ACTUAL FOR AUSTRALIA	28
FIGURE 3.7. LSTM WIND POWER PREDICTION FOR AUSTRALIA	28
FIGURE 3.8. SOLAR POWER GENERATION ACTUAL FOR AUSTRALIA	29
FIGURE 3.9. LSTM SOLAR POWER PREDICTION FOR AUSTRALIA.....	29
FIGURE 3.10. WIND POWER GENERATION ACTUAL FOR NETHERLANDS	30
FIGURE 3.11. LSTM WIND POWER PREDICTION FOR NETHERLANDS.....	30
FIGURE 3.12. SOLAR POWER GENERATION ACTUAL FOR NETHERLANDS.....	31
FIGURE 3.13. LSTM SOLAR POWER PREDICTION FOR NETHERLANDS	31
FIGURE 3.14. MEAN SQUARE ERROR RESULTS	33
FIGURE 3.15. ROOT MEAN SQUARE ERRORRESULTS.....	33
FIGURE 4.1. MASTER / SLAVE ARCHITECTURE IN HADOOP.	40

List of tables

TABLE 1: COUNTRY, ATTRIBUTES, OF DATA	26
TABLE 2: NODE1 AUSTRALIA WIND ENERGY FORECASTING.....	32
TABLE 3: NODE1 AUSTRALIA SOLAR ENERGY FORECASTING.....	32
TABLE 4: NODE2 NETHERLANDS WIND ENERGY FORECASTING	32
TABLE5: NODE2 NETHERLANDSSOLAR ENERGY FORECASTI	33

General Introduction

General Introduction

General Introduction

Renewable energy is an important research field. It concern the energy of wind tribune or solar panel and the other sources of renewable energy. Energy forecasting studies the variation of the energy generation, consumption or energy optimization.

The growing of smart renewable energy take the attention of several works [1,2,3,4, 5,7,8]. These studies proposed several paradigms to renewable energy forecasting, statistical such as ARIMA Model, GARCH Model [6] or simple machine learning models such as Support Vector Machine, linear regression k-means ...etc. Although these models give a results but it stayed limited.

Deep learning model is introduced to analyze time series. The LSTM network (Long-Short Term Memory) become a good deep learning model to process and predict important events with long intervals and delays in the time series. We called to LSTM kernel to forecasting renewable energy.

We presented our work in four chapters, the first chapter a presentation on renewable energies and their types, we also focused on the solar panels and wind tribunes that are the subject of our research.

The second chapter concern the concepts of deep learning, machine learning, long short-term memory as we talked about the big data and internet of things.

we presented in Chapter three the research proposition, which is an architecture for predicting energy, and related works, with the results obtained.

In the final chapter we showed the programming environment and implementation with validation of results.

Chapter 01: State of the art: renewable energy

Chapter 01:
State of the art: renewable energy

Chapter 01: State of the art: renewable energy

1 Introduction

Countries, Societies and companies are today racing to have the most mixed renewable energy resources. This racing started because of different problems posed by using unrenovable energy resources such cost increasement, global warm and many other problems like pollution and so on.

In this chapter we are going to define energy and its type passing to its types renewable and unrenovable. Walking through those types we are going to site each type and focusing on the principal ones telling their advantages and components.

2 Definition

First, we need to define energy before we represent its types, so duo to **oxford dictionaries** energy can be defined as: “The strength and vitality required for sustained physical or mental activity.” Or: “Power derived from the utilization of physical or chemical resources, especially to provide light and heat or to work machines.”. From that point we continue the definition of energy by stating its types: it exists two types of energy that have common points such that they are both natural resources but the first is limited, and we name that one **Non-renewable** energy. The second is the opposite unlimited resource, we name it **renewable** energy.

Non-renewable energy cannot be re-made or enchase or regrowth in short term. **Renewable** energy in the other hand, can be regenerated in a short time or in a short period. [9]

3 Renewable energy types

Generally Renewable energy is generated from natural resources which are renewable and can be generated in a short time such as sunlight, wind, rain, tides and geothermal heat.

- Solar energy
- Wind energy
- Hydropower
- Biomass
- Ocean energy
- Geothermal energy
- Waste to Energy
- Cultivated Plants

So, we are going to concentrate in the two first ones because they are our study subject.

3.1 Solar Energy

The energy that the sun sends to our planet earth in one hour, is the same amount of energy that the world population consume in one year. So, in addition solar energy is considered as the cleanest energy source. Still we need to say what is solar power, simply it's transforming the sunlight into a consumable energy such as electricity. That transformation

Chapter 01: State of the art: renewable energy

can be done using two ways **photovoltaics** (figure 1) and it's done directly or **concentrated solar power** (figure 2). [9]



FIGURE 1.1. concentrated solar power



FIGURE 1.2. Photovoltaics

3.1.1 Solar energy component

Solar power simple system can be composed as shown in (Figure 3). [10]

- **Energy Source:** Of course, in our case the energy source is the sunlight or the sun heat.
- **Energy Conversion:** Photovoltaics (PV) panels Their placement and mounting affect your system performance more than any other facet of the job.
- **Energy Inversion:** Inverters take the low-voltage, high-current signals from the PV panels and convert them into 120VAC (or 240 VAC), which is directly compatible with grid power.

Chapter 01: State of the art: renewable energy

- **Energy Storage:** A device that stores energy is generally called an accumulator or battery. Must be a rechargeable battery.
- **Energy Distribution:** its job is the delivery of electric power; it takes electricity from the transmission system to used energy or to the grid (electric utility). Distribution substations connect to the transmission system to lower the transmission voltage to medium voltage using transformers so it be in a range from 2 kV to 35 kV
- **Energy use:** Houses, Cars
- **Electric Utility:** Other uses

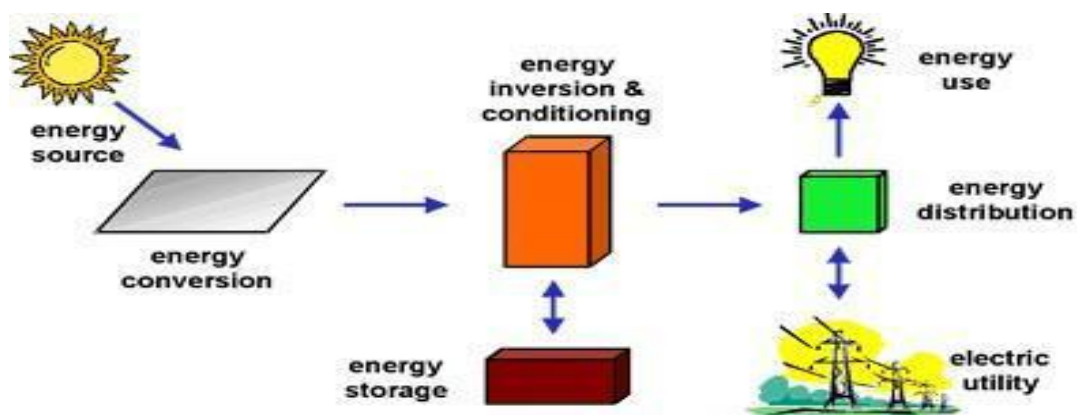


FIGURE 1.3. Solar System

3.1.2 Advantages

- Solar power helps to slow/stop global warming.
- Solar power saves money.
- Solar power provides energy reliability.
- Solar power provides energy security.
- Solar power provides energy independence.
- Solar power creates jobs.

3.2 Wind energy

One of the first energy resources harvested by us human was wind energy. We used wind to push our sailboats, and helped mills to grind grains or pump water. In principle wind is an air that is generated from the temperature changing so the air moves from high pressure to low pressure.[10]

Chapter 01: State of the art: renewable energy



FIGURE 1.4. Wind energy

3.2.1 Wind energy component

- **rotor blades:** Capturing wind energy then it has to convert it into energy (rotational energy)
- **shaft:** Transferring the produced energy from the rotor blades to the generator.
- **Nacelle:** is the cover or the envelope that contains the **gearbox**
- **Gearbox:** increasing the speed of the shaft between two components rotor hub & generator.
- **Generator:** generating electricity from the energy of the shaft and it uses electromagnetism.
- **Electronic control unit:** responsible of shutting down turbine if the system is malfunctioning also monitors the system and controls yaw mechanism.
- **Yaw controller:** keep track of wind alignment by moving the rotor towards it.
- **Brakes:** stop the shaft in case of power overload or system failure.
- **Tower:** lifts the entire setup.

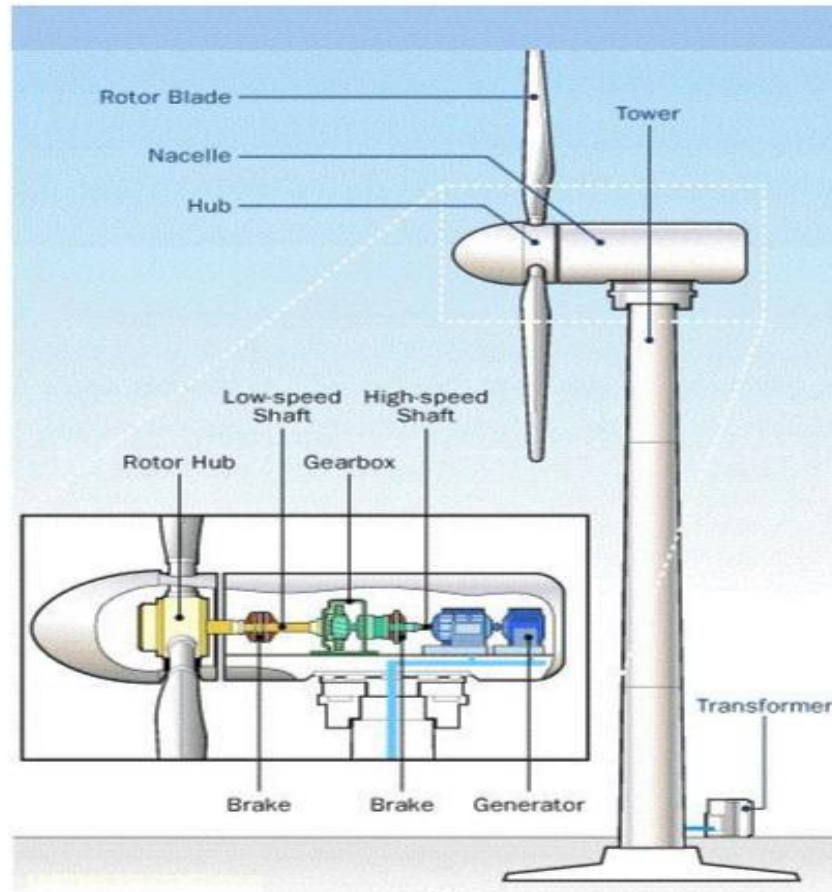


FIGURE 1.5. horizontal-axis wind turbine

3.2.2 Advantages

- Clean, mean no emission
- Conserves air quality and water quality
- Help in climate changes
- Domestic energy
- Energy independence

4 Renewable energy comparison

Comparing renewable energy technologies solar power versus wind power is not that simple or easy because each of those technologies has their advantages as we saw in the previous section. It usually depends on the application location of the system.

The first obvious difference between the two is wind power used to aliment big utilities such as national grid application so the wind power is way more effective than solar power in big scale projects and commercial projects. We may ask why? The answer is most of time wind turbines never stop generating energy day and night all the day, but in the other side solar panels can generates energy only in day light when the sun existing so it can generate any energy if any inappropriate weather changes happened. Secondly wind turbines can be installed in all locations even far away in the sea. But solar panels are confined in some places only.

Chapter 01: State of the art: renewable energy

The second difference that we can locate is solar panels can be installed in tight strict places and wind turbines is the total opposite it needs large open spaces and most times it needs to be in high places. Solar panels also has better reliability, and based on data from NASA and BOM it provides as an output a predictable energy. [9][10]

5 Why use Renewable energy

The world today faces so many problems one of the major and most critic problems is energy duo to its important and the increasing cost of extracting energy and the day after day those resources are decreasing and have bad emission that is costing global warm issue.

Renewable energy resources such as solar photovoltaic (PV), concentrated solar power (CSP), wind power, and biomass have been growing rapidly in energy market. Even those resources are endless but there are other facts that can limit them which can cause instability in producing energy. This is because those resources are related to environmental changes so the sun can be blocked by the clouds and the sun only available on the day, the winds sometimes are there other times aren't there. Overcoming that problem is easy and the solution is by storing the energy, yet the storing process is costly and some problems like dividing the energy some need to be stored and other need to be stored for later.

Those problems are counted above awakened the challenge for Artificial Intelligence to work in solving those problems, and provide a better energy management generation and consumption and storage.

Machine learning is used by its different techniques for renewable energy different problems in different stages. [11]

Chapter 01: State of the art: renewable energy

6 Conclusion

In this chapter, we tried to clarify the uses of renewable energy and some of its types and we compared between the main ones and gone far making a reasonable and appropriate causes of way using renewable energy and why uses artificial intelligence to solve problems posed in that domain

In the next chapter, we are going to clarify and talk about machine learning techniques and deep learning and other technologies. Reaching the main method and technique that we are going to use in our work.

Chapter 02: Deep learning and big data system

Chapter 02: Deep learning and big data system

Chapter 02: Deep learning and big data system

1 Introduction

Artificial intelligence and his different approaches are on its golden age. When we can see machine learning adapting and developing so fast given such precise results on different domains. Machine learning most important approaches is also living on the peak, deep learning develops the hardware and the processing power which allows to applicate the algorithms and work on a huge amount of data without facing any troubles.

In this chapter, we are going to talk about machine learning and deep learning then the algorithms of the deep learning going through the neural network's architecture passing after that to LSTM the method we are going to use on solving the problem of forecasting of renewable energy.

2 Deep learning

2.1 Definition

Deep learning is a branch of machine learning and the last is a part of artificial intelligence. Deep learning is based completely on artificial neurons, and it duplicates the work of neural network in human brain so deep learning is also aiming on duplicating the human brain and the way it works. So, in deep learning we are not obligated to program everything.

Deep learning it has been around for awhile so it is not a new concept, nowadays it is on its gold era because in the past we didn't have much processing power and that much amount of data.[12]

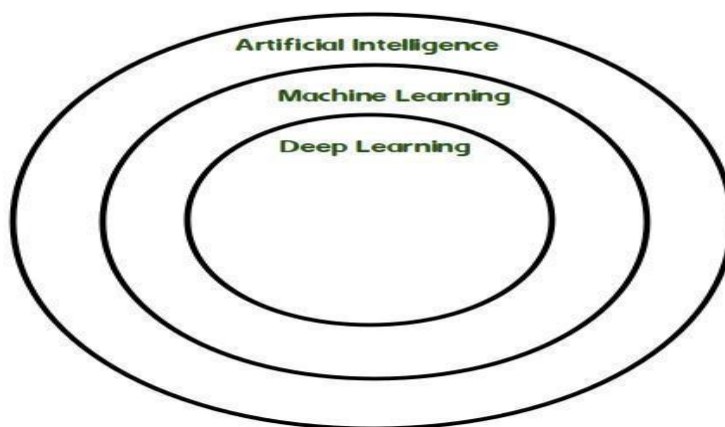


FIGURE 2.1. Deep Learning & Machine Learning

2.2 Architecture

Defining deep learning as a neural network poses a question what kind of neural network is it? It can be referred as classes of neural networks with many hidden layers called Deep Neural Network DNN. DNN applications speech and image recognition,

Chapter 02: Deep learning and big data system

machine translation, and forecasting of financial time series. The four fundamental network architectures:

- **Unsupervised Pre-trained Networks:** It has multiple layers and it is able of processing non-linear relations, also it is complex for a certain level
- **Convolutional Neural Networks:** It is considered as standard neural network; CNN is extended across space using shared weights.
- **Recurrent Neural Networks:** This neural network has edges which supply the timepoint in state of the next layer of the same time step. It is intended to perceive sequences, such as speeches and texts. It also has a short memory as cycles in the net.
- **Recursive Neural Network:** The network is not forced to work in a time aspect so that the input goes in a hierarchically network as a tree.

It does exist other architectures such as **Deep Belief Network DBN**. Deep learning is an active research field and it has multiple applications in the domain of renewable energy forecasting and those applications and studies are focused on wind and solar power. Those applications used different deep learning algorithms such as LSTM, DBN, ... etc.

3 Deep learning Algorithms

In this section, we are going to introduce some of deep learning algorithms such as back propagation and long short-term memory (LSTM) and Dropout.

3.1 Back propagation

Back-prop is one of the various deep learning algorithms. It used for regression problems and classifications also it used to train Multi-Layer Perceptron (MLP), MLP is a basic architecture contains multiple connected layers of neurons. Back-prop is used to compute the gradient of a function as an example gradient descent algorithm that solve optimization problems using the back propagation. Computing the gradient can be done in two famous ways: the first **Analytic differentiation** in this situation we just calculate derivatives by the chain rule.

The second way is **Approximate differentiation using finite difference** this one has big complexity compared to the first one. The complexity is estimated as $O(N)$ that is because of the evaluation function. N considered as number of parameters. [14]

3.2 Dropout

The most powerful machine learning systems are deep neural nets with big number of parameters, anyway that's pose a problem in how fit later on. Another problem is the speed because large networks are not that fast making test time so long and combining predictions to deal with overfitting.

However, the algorithm or technique we are going to discuss now handling the problems we posed. [15]

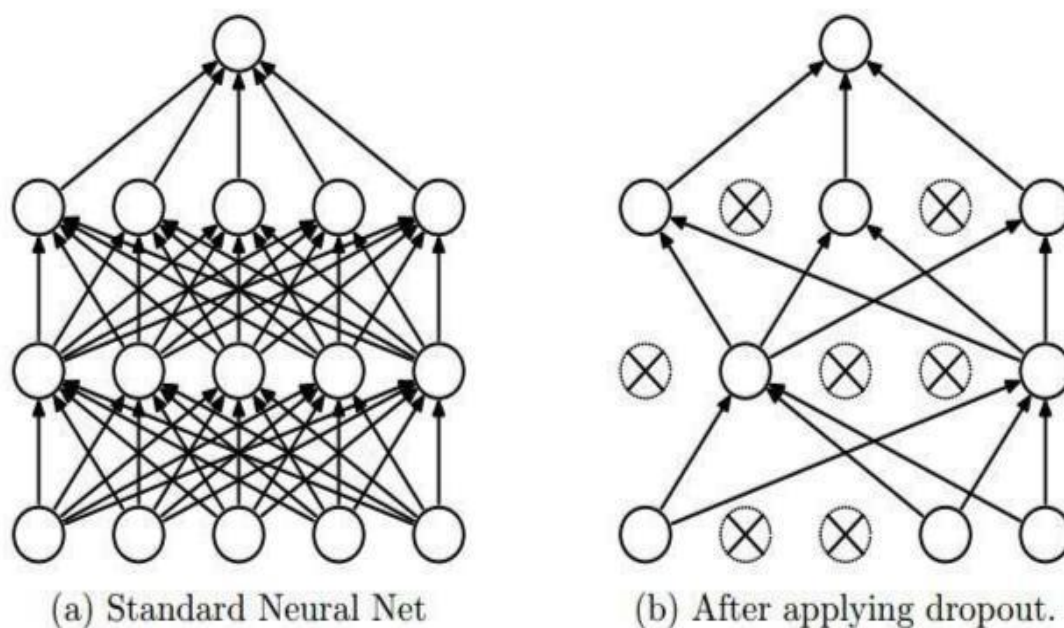


FIGURE 2.2. Drop Out

It was introduced by Hinton et al. (2012). The concept of the drop out neurons with their connections randomly in the training phase. This method improves remarkably the generalization properties of deep neural networks. Any way as any other it has some disadvantages such slowing the training time.

3.3 Long Short-Term Memory (LSTM)

LSTM is our chosen method in solving our problem. A LSTM network is based on RNN but it has some different aspects that differentiate it from other neuron nets in a recurrent neural network. The first is the LSTM has control on the input, meaning it can decide when the neuron can receive input and when it can't.

The second aspect it has the ability to decide when to remember what is the last or previous operation. The third it has control on passing the output to the next step or not. Those are some basic aspects and options in LSTM in the next section we are going to talk about it more in details. [14][15]

3.3.1 LSTM characteristics

In the last few years, RNN has achieved great results in different applications such as speech recognition and images also translation, the success achieved can be assigned to Long Short-Term Memory (LSMT). Which is a special type of RNN, was introduced by Hochreiter and Schmidhuber(1997). The special about LSTM that is unlike the ordinary MLP RNN and has a special neurons that has a memory that structure called Memory cell.

Those cells has the ability to store information for an amount of time. The LSTM difference from ordinary RNN appear also in the ability of LSTM to control and decide on, when to the input enter or not, to remember the previous result or not and when the output goes to the next step. Each of what mentioned (input, operations, output) has its own activation function,

Chapter 02: Deep learning and big data system

but what is really wonder full is the LSTM decides on the current input. [13][14][15]

3.3.2 LSTM compositions

LSTM is composed of three gates to control the flow of the data entering to the neuron and leaving. Each gate has its own activation function and the neuron in LSTM has the same input as in normal neurons. [13]

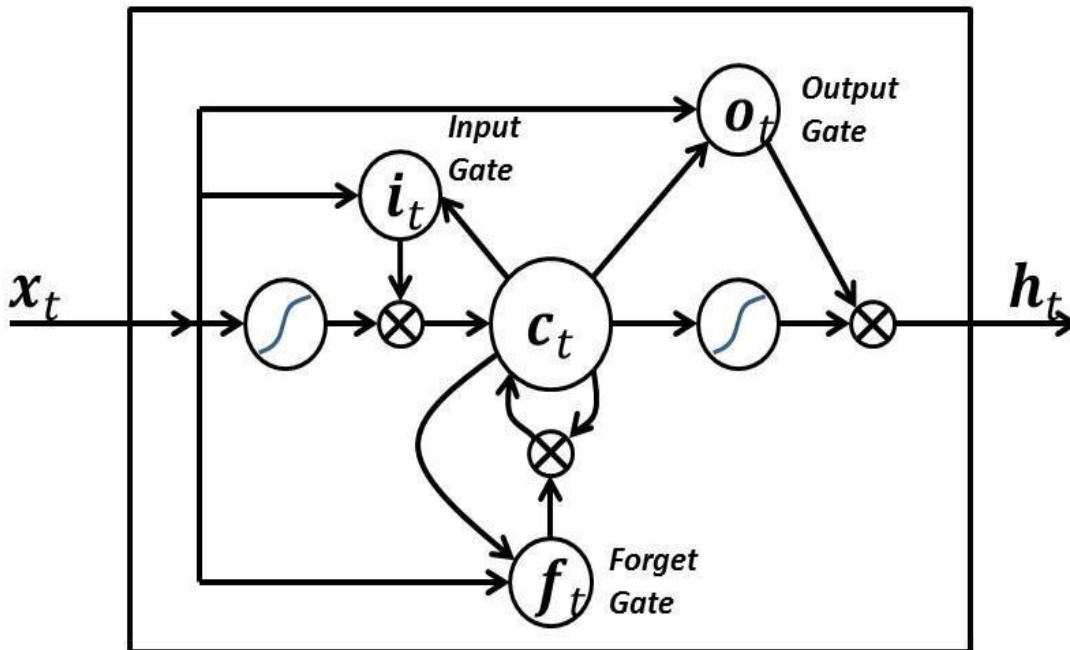


FIGURE 2.3. LSTM structure [31,32]

The principal of operation is so simple the input signal X_t at the current time t has the decision of all the other gates, for example: the input gate it gets a high signal, then the input gate it stores the signal. If the output gate O_t gets high signal it releases the stored data to the next neuron. If the forget gate F_t gets a high signal the memory cell get cleared. So, ordering who takes the decision first, the input gats come first then the forget gate finally the output gate. However, the only gate that is capable of taking all of the three decision is the input gate. That order and way of work simulates the way our brains work.

Chapter 02: Deep learning and big data system

4 Internet of Things IoT

The emergence of IoT come from the fast growing of internet and the changing of its shape from a traditional content of texts, images and videos to physical objects that is controlled by users.

The term was first used by Kevin Ashton (MIT) in 1999. The use of IoT can be seen surveillance camera, cars, weather stations and so on. Smart energy frameworks also uses IoT to automate and control energy between buildings. [16][17]

Internet of Things (IoT), is defined by the ITU (International Telecommunication Union) and IREC (IoT European Research Cluster) “as a dynamic global network infrastructure with self-configuring capabilities. Based on standard and interoperable communication protocols where physical and virtual "things" have identities. Physical attributes and virtual personalities, use intelligent interfaces and are seamlessly integrated into the information network.” [16][17]

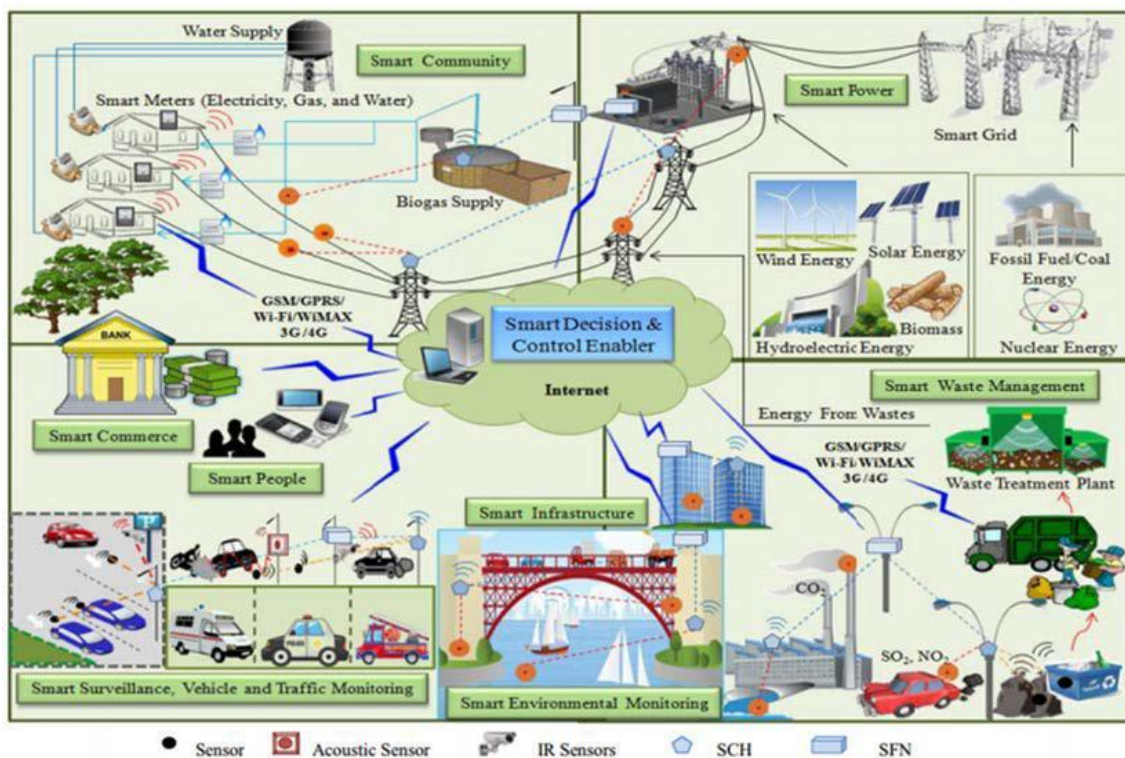


FIGURE 2.4. Area of Internet of Things

Chapter 02: Deep learning and big data system

4.1 Area of application [16][17]

- Agriculture: IoT allow us more advanced remote control of farms, resulting a better sensing and monitoring of production.
- Education: The integration of IoT make the education more interactive, and students be actively engaged which was proven scientifically a better way to learn.
- Energy: IoT giving us the ability to economize the energy, by offering the power of decisions making to the thing's "objects", so they be able to decide whenever is a good time to turn on or off.

It does exist other area of applications such as Cars, Health Care and smart cities ... etc.

4.2 Platforms [16][17]

4.2.1 Wearable platform

Wearable platform is gaining popularity each day duo to the ability of using it in health care, daily business and the emerge off pebble watch that works with both apple and android allowing us to know our health and heart beats and so many other things.

4.2.2 Embedded platform

Embedded software solutions can be used to connect different devices that meet the reliability, security and power consumption, the demand on those platforms is on all industries such medical and industrial even automotive. The best starting point probably is Arduino, and other advanced such Raspberry Bi.

Chapter 02: Deep learning and big data system

5 Big Data

Big data appeared and has been known the last few years, and it's one of the most discussed topics in the market today in all industry sections. The word "Big Data" refers to large dataset collected by governments or societies or companies.

The difference between normal and big data is that big data so large and complex that we can't process with normal data processing methods also it is so valuable because of the variety of data it contains it and the hidden data structure. That term was founded by Doug Laney in 2000 and the definition of big data is the three V (Volume, Velocity and Variety) [18][19]

1. Volume: The size of data that is gathered by companies or government from agencies or users the data size usually is measured by terabytes or zettabytes. In the early starts the problem faced was how to store that huge amount of data but today the problem is solved and the new problem is about computing those data. [18][19]

2. Velocity: is the speed that data are collected. Because today data is not lagged instead it is collected in real time at a very enormous speed. [18][19]

3. Variety: The different types of data collected. When we say different, we mean maybe same data type but different sources, or different types but the same source or it can be both different sources and types of data. More than that data can come in different format meaning database tables or semi-structured data such as XML data, unstructured data such as text, images, video streams, audio statement, and more. [18][19]

Another concept lately was added is veracity. This concept specialized in defining which data are important and which are not to make the work on those data more useful and precise. [19]



FIGURE 2.5. Three V for big data

Chapter 02: Deep learning and big data system

7 Conclusion

In this chapter, we talked about artificial intelligence and deep learning. We did dive into deep learning defining it and counting some of its algorithms then we passed to neural network and the architecture of neural networks.

We also talked about algorithms of deep learning such as back-prop and drop out then LSTM in details because it's our chosen method to solve the problem of forecasting of renewable energy.

In the next chapter, we are going to show our architecture and system conception and how we used the LSTM to solve the problem and some similar systems that has used the same solution.

Chapter 03: proposition of deep learning for renewable energy

Chapter 03:
Proposition of deep learning for
renewable energy

Chapter 03: proposition of deep learning for renewable energy

1. Data extraction module: Responsible for data acquisition, the user can use Internet of things devices in this module in order to captured data.
2. Data storage module: Concern the storage system (storing and processing) of data gathered from previous module, like a HADOOP FILE SYSTEM (HDFS).
3. Machine Learning Module: This module can analysis data gathered in Data storage module in order to create prediction model for energy forecasting.

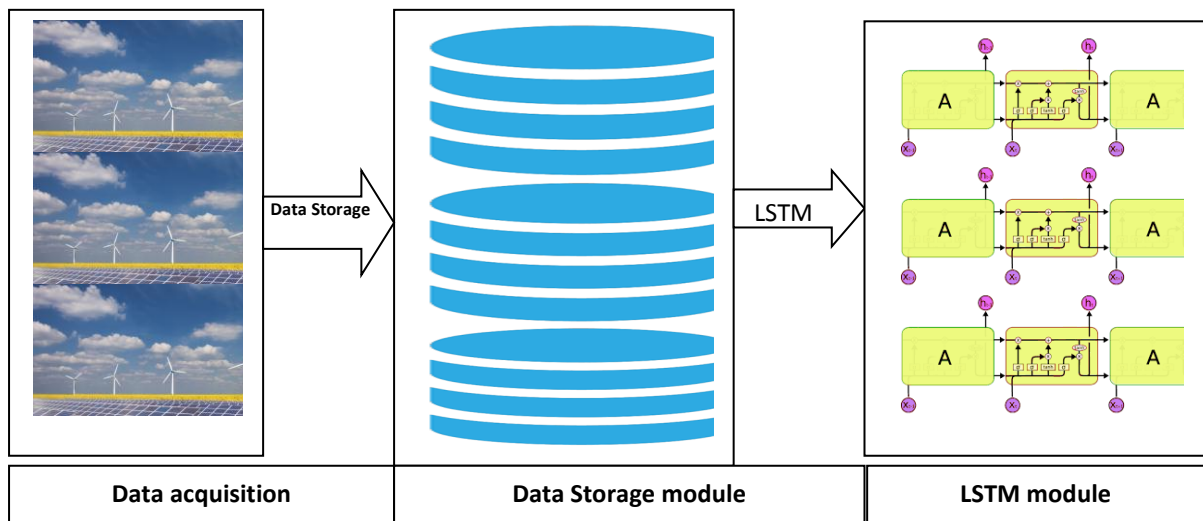


FIGURE 3. 2. Architecture of our system

Chapter 03: proposition of deep learning for renewable energy

3.2 Data Mining Process For Renewable Energy

Data mining process is a multiple step follow by the developer of system to implement it. Cross-industry standard process for data mining called CRISP-DM [37,39] is a standardization for data mining process.

We propose CRISP-DM for our system This process enable developers to implement renewable energy software easily, as well as display the source of energy.

CRISP-DM: Phases are:

- **Business Understanding:** Understanding renewable energy objectives and requirements.
- **Data Understanding:** Initial data collection from solar panel or wind tribune or other renewable energy sources.
- **Data Preparation:** Record and attribute selection; Data cleansing.
- **Modeling:** In this phase we apply LSTM model.
- **Evaluation:** Determine if results meet renewable energy forecasting; Identify business issues that should have been addressed earlier.
- **Deployment:** Put the resulting models into practice; Set up for continuous mining of the data

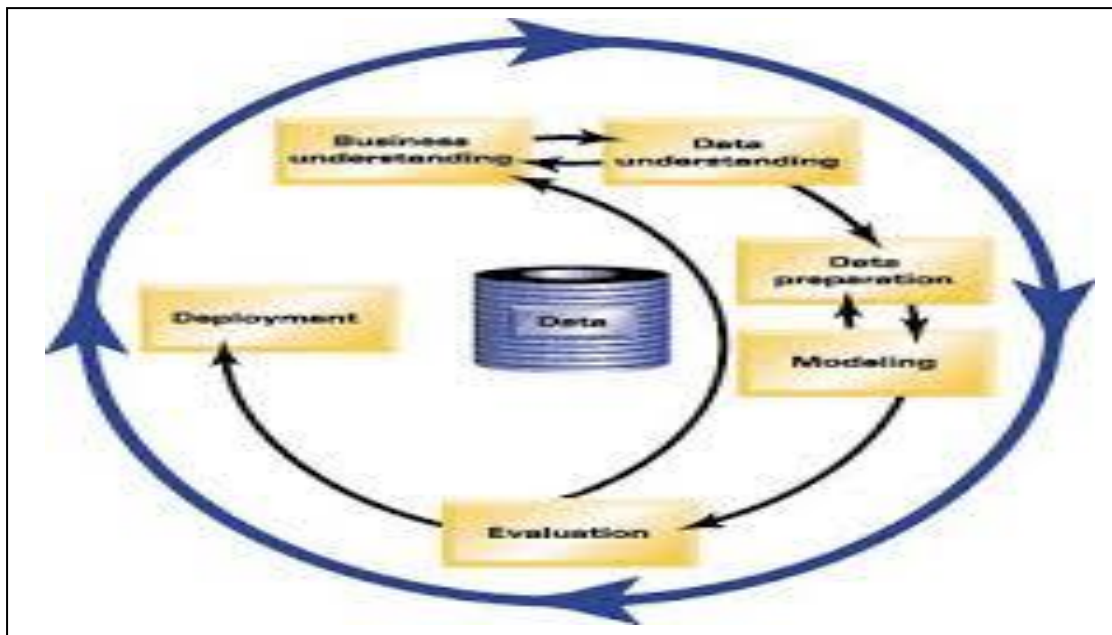


FIGURE 3. 3. CRISP-DM

3.3 Long Short-Term Memory Algorithm

We propose LSTM deep learning algorithm for renewable energy forecasting, this algorithm has the same kernel of LSTM algorithm.

LSTM network is a type of recurrent networks. it introduced to treat the problem that have time dimension, in which Recurrent neuronal networks cannot track the relationship across long time. It's specifically proposed to learn long term dependencies and able to resolve vanishing and exploding problems exist in recurrent neural network [30].

By adding a set of memory units to Recurrent neural networks, LSTM effectively solve the vanished gradient problem or gradient explosion.

The components of LSTM are three gates for each unit which are forget gate input gate and output gate. Memory cell records all historical information in the current time (t) from the three gates forget gate (ft) input gate (it) and the output gates (ot). The values of gates between [0..1].

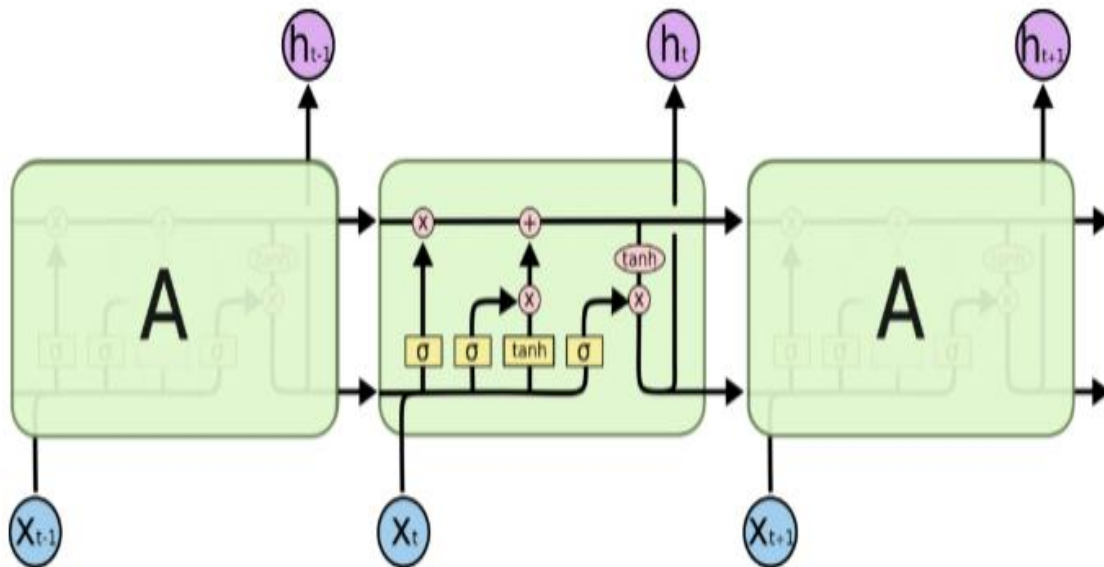


FIGURE3. 4. LSTM cells connected to each other

Equations bellow for the forget input (f_t), input gate (i_t) and output gate (o_t):

$$f_t : \text{Forget gate activation: } f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f)$$

$$i_t : \text{Input gate activation: } i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i)$$

$$o_t : \text{Output vector of LSTM: } o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o)$$

Chapter 03: proposition of deep learning for renewable energy

5 Conclusion

In this chapter we proposed an architecture for predicting renewable energy, It allows to easily collect data from data acquisition module of renewable energy to the, and stored through a second module of the system, In the Machine learning module we propose deep learning .We found that the machine learning of Long Short Term Memory (LSTM) algorithm results are much better than the other machine learning algorithms: Linear Regression (LR), K-nearest neighbors (KNN) and Classification And Regression Trees (CART).

Chapter 04: implementation and development

**Chapter 04:
implementation and development**

Chapter 04: implementation and development

1 Introduction

The problematic discussed in the third chapter can be classified as one of the problems of predicting the time series which is considered as a difficult type of predictive modeling problem.

We can consider the algorithm Long-term memory distributed or LSTM is one of the types of Recurrent neural networks used in deep learning, also the recurrent neural networks are called a strong type of neural network designed to deal with sequence dependency, Because the model building can be trained successfully mostly.

So the accuracy of Training, validation and performance are all one of several ways for a good evaluation of the neural network it is the amount of learning and effectiveness of the results.

2 Used tools

2.1 Language of Python

Python is a powerful fast and easy to learn, is an open source programming language and it an interpreted high-level programming language and object-oriented programming, developed by Python Software Foundation, it implementation began in December 1989.[40]

It's easy to work with it and have access to a high-quality library, It is the most used in the area of artificial intelligence and Machine learning. The most famous libraries in the field of artificial intelligence are made by python like TensorFlow and Scikit-learn, Keras, TensorFlow guides to use Python language; although there are other programming languages use this library, but it's not as effective as Python. [41]



Chapter 04: implementation and development


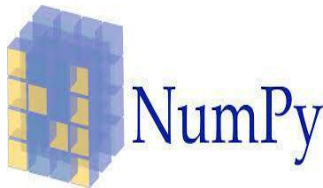
2.2 Jupyter notebook

The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more, The Notebook has support for over 40 programming languages, including Python.[42]






2.3 Libraries used

We have used a set of libraries as needed in this implementation, each one played its role according to its competence, as shown in the following table:

Library	Logo	Definition
Pandas		Pandas is a library written for the Python programming language for manipulating and analyzing data. The main data structures are the series (to store data according to a dimension - size according to an index), the DataFrames (to store data according to 2 dimensions - rows and columns), the Panels (to represent data according to 3 dimensions, Panels4D or DataFrames with hierarchical indexes also named MultiIndex (to represent data in more than 3 dimensions - hypercube)[43]
NumPy		NumPy is a library for the Python programming language, adding support for large, multidimensional arrays and matrices, along with a large Collection of high level mathematical functions to operate on these arrays. NumPy is open-source software and has many contributors.[44]
Matplotlib		Matplotlib is a library of the Python programming language designed to plot and visualize data in graphical form. It can be

Chapter 04: implementation and development

		<p>combined with the NumPy and SciPy scientific python libraries. [45]</p> <p>Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and IPython shells, the Jupyter notebook, web application servers, and four graphical user interface toolkits.</p> <p>You can generate plots, histograms, power spectra, bar charts, errorcharts, scatterplots, etc., with just a few lines of code.[46]</p>
Keras		<p>Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano.</p> <p>Supports both convolutional networks and recurrent networks, as well as combinations of the two.[47]</p> <p>Keras is an open source neural network library. Designed and Allows to enable fast prototyping and experimentation with deep neural networks, it focuses on being user-friendly, modular, and extensible.</p> <p>Chollet explained that Keras was conceived to be an interface rather than a standalone machine-learning framework. It offers a higher-level, more intuitive set of abstractions that make it easy to develop deep learning models regardless of the computational backend used.[48]</p>
sklearn		<p>Scikit-learn is a free Python library for machine learning.</p> <p>It includes functions for estimating random forests, logistic regressions, classification algorithms, and support vector machines. It is designed to harmonize with other free Python libraries, including NumPy and SciPy.[49]</p>

Chapter 04: implementation and development

2.4 Hadoop big Data tool's

The system of Hadoop File System (HDFS) is characterized by the storage of these files on a large number of machines so as to make invisible the exact position of a file with the transparency of access.

It is also characterized by copying files in multiple copies for reliability and allowing multiple simultaneous accesses.

HDFS lets you see all the folders and files of these thousands of machines as a single tree containing them as if they were on the local hard drive.[50]

The HDFS do exchange with world with two commands like example :

- to put file in HDFS :`hdfsdfs -copy FromLocalfichiersrfichierdst .`
- to extract file from HDFS :`hdfsdfs -copyToLocalfichiersrcdst.`

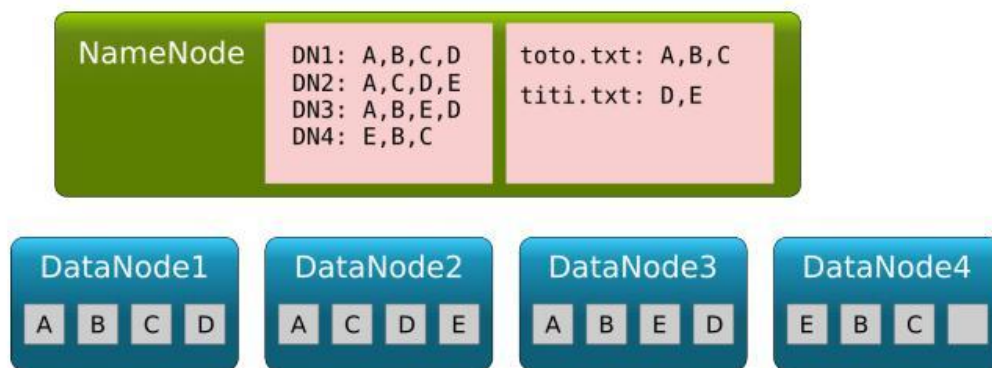


FIGURE 4.1. Master / Slave Architecture in Hadoop. [50]

An HDFS cluster consists of machines playing different exclusive roles among them:

- One of the machines is the HDFS master, called the namenode. This machine contains all the names and blocks of files, like a big phone book.
- Another machine is the secondary namenode, a kind of backup namenode, which records backups of the directory at regular intervals.
- Some machines are customers. These are access points to the cluster to connect to and work.
- All other machines are datanodes. They store the blocks of the contents of the files.



Chapter 04: implementation and development

3 Programming Environment

3.1 Loading of Data set

We use the data of two countries with CSV format Austria (AT) and Netherlands (NL). Data contain the information about renewable energy solar and wind energy in hourly resolution. the using of Pandas library can make load this dataset easily.

***** Australia *****

```
In [1]: import pandas as pd
production = pd.read_csv("/home/dell/expt/data/time_series_60min_singleindex.csv",
                        usecols=(lambda s: s.startswith('utc') | s.startswith('AT')),
                        parse_dates=[0], index_col=0)
```

```
In [2]: production.head(3)
```

Out[2]:

utc_timestamp	AT_load_entsoe_power_statistics	AT_load_entsoe_transparency	AT_solar_generation_actual	AT_wind_onshore_generation_actual
2005-12-31 23:00:00+00:00	6297.0	NaN	NaN	NaN
2006-01-01 00:00:00+00:00	6005.0	NaN	NaN	NaN
2006-01-01 01:00:00+00:00	5743.0	NaN	NaN	NaN

```
In [3]: production.tail(3)
```

Out[3]:

utc_timestamp	AT_load_entsoe_power_statistics	AT_load_entsoe_transparency	AT_solar_generation_actual	AT_wind_onshore_generation_actual
2018-05-31 21:00:00+00:00	NaN	5113.0	0.0	303.0
2018-05-31 22:00:00+00:00	NaN	5167.0	0.0	251.0

***** Netherlands *****

```
In [50]: #read Netherlands dataset
import pandas as pd
production = pd.read_csv("/home/dell/expt/data/time_series_60min_singleindex.csv",
                        usecols=(lambda s: s.startswith('utc') | s.startswith('NL')),
                        parse_dates=[0], index_col=0)
```

```
In [51]: import numpy as np
def clean_dataset(df):
    assert isinstance(df, pd.DataFrame), "df needs to be a pd.DataFrame"
    df.dropna(inplace=True)
    indices_to_keep = ~df.isin([np.nan, np.inf, -np.inf]).any(1)
    return df[indices_to_keep].astype(np.float64)
```

```
In [52]: clean_dataset(production)
```

Out[52]:

utc_timestamp	NL_load_entsoe_power_statistics	NL_load_entsoe_transparency	NL_solar_generation_actual	NL_wind_offshore_generation_actual	NL_wind_onshore
2015-01-01 00:00:00+00:00	11338.0	9456.0	0.0	145.0	
2015-01-01 01:00:00+00:00	10917.0	8957.0	0.0	145.0	
2015-01-01 02:00:00+00:00	10473.0	8462.0	0.0	148.0	
2015-01-01 03:00:00+00:00	10200.0	8156.0	0.0	134.0	

Chapter 04: implementation and development

We can observe that the production of renewable energy by wind is often available during the year and this is come back to wind currents in these countries.

Either the production of renewable energy through the sun in the two countries Linked to peak times especially in the summer.

However, renewable energy production by wind is more than double or triple Compared to the sun.

3.3Elaboration of Data set

We used all the import functions and classes. with a SciPy environment and install Keras deep learning library.

So to modeling our neural network we need to convert the integer values to floating point values, after extract the NumPy array.

when we activate the sigmoid function, we must rescale the data to the range of 0-to-1 with using the MinMaxScaler from sklearn library.

```
In [25]: from keras.models import Sequential
         from keras.layers import Dense
         from keras.layers import LSTM
```

Using TensorFlow backend.

```
In [26]: from sklearn.preprocessing import MinMaxScaler
         from sklearn.metrics import mean_squared_error
         import numpy
         import matplotlib.pyplot as plt
         import pandas
         import math
         # fix random seed for reproducibility
         numpy.random.seed(7)
```

```
In [27]: # convert the integer values to floating values "more suitable for modeling NN"
         dataset=wind.values
         dataset = dataset.astype('float32')
         dataset=dataset.reshape(-1,1)
         print(dataset)
```

```
[[152.]
 [110.]
 [ 56.]
 ...
 [114.]
 [ 94.]
 [ 93.]]
```

```
In [28]: # normalize the dataset:LSTMs(sigmoid) sensitive VS data / it muste to rescale whith MinMaxScaler(0,1)
         scaler = MinMaxScaler(feature_range=(0, 1))
         dataset = scaler.fit_transform(dataset)
```

```
In [29]: print(dataset)
```

```
[[0.05963122]
 [0.04315418]
 [0.0219694 ]
 ...
 [0.04472342]
 [0.03687721]
 [0.03648489]]
```

Chapter 04: implementation and development

3.4 Train and Test Data Set

the sequence of values in time series data get us a simple method to split our dataset into train and test datasets, we chose separate it in the split point 67% for training datasets and 33% for testing the model.

```
In [30]: # split into train and test sets
train_size = int(len(dataset) * 0.67)
test_size = len(dataset) - train_size
train, test = dataset[0:train_size:], dataset[train_size:len(dataset):]
print(len(train), len(test))

5885 2899
```

after indexing one of our dataset like wind or solar in Node 1 or wind or solar in Node 2 ,we need to convert our single column of data into a two column dataset: one containing energy product in this hour (t) and the second column containing next energy product in hour (t+1) , to be predicted. to make it we made a function takes two arguments:

- * the dataset(convert a NumPy array to a dataset)
- * the look_back (input variables such as previous time steps -in this case 3- use to predict the period of the next time), so we will create a dataset where X is the amount of energy at a given time (t) and Y is the number of amount of energy at the next time (t +1).

```
In [31]: # convert an array of values into a dataset matrix
def create_dataset(dataset, look_back=1):
    dataX, dataY = [], []
    for i in range(len(dataset)-look_back-1):
        a = dataset[i:(i+look_back), 0]
        dataX.append(a)
        dataY.append(dataset[i + look_back, 0])
    return numpy.array(dataX), numpy.array(dataY)
```

```
In [32]: # where X is power wind in time (t) and Y is power wind in the next time (t + 1).
look_back = 3
trainX, trainY = create_dataset(train, look_back)
testX, testY = create_dataset(test, look_back)
```

Our data is in the form: [samples, features] and because our LSTM network is sensitive of the input data (X) to be provided with a specific array structure in the form of: [samples, time steps, features].we will by constructing a differently shaped dataset by framing the problem as one time step for each sample.

So we used numpy.reshape() to prepared expected structure through the transform of train and test input data.

Chapter 04: implementation and development

```
In [34]: # reshape input to be [samples, time steps, features] (LSTM expects input data (X) with a specific array str
trainX = numpy.reshape(trainX, (trainX.shape[0], 1, trainX.shape[1]))
testX = numpy.reshape(testX, (testX.shape[0], 1, testX.shape[1]))
```

```
In [35]: print(trainX)
```

```
[[[0.05963122 0.04315418 0.0219694 ]]
 [[0.04315418 0.0219694 0.01530012]]
 [[0.0219694 0.01530012 0.01333856]]
 ...
 [[0.04550804 0.04864653 0.06786975]]
 [[0.04864653 0.06786975 0.06904668]]
 [[0.06786975 0.06904668 0.06630051]]]
```

3.5LSTM : create, fit and predictions

LSTM network that we are design has the follow proprieties:

- 1-one visible layer with 1 input .
- 2-a hidden layer with 4 LSTM blocks .
- 3-an output layer that makes a single value prediction.
- 4-default sigmoid activation function.
- 5-trained for 10 epochs and a batch size = 1.
- 6-use mean_squared_error for loss gate and adam's optimizer for propagation back.

The estimate performance of our model (after training) on the dataset (train and test) give us a reference for next models.

```
In [36]: # create and fit the LSTM network
model = Sequential()
model.add(LSTM(4, input_shape=(1,look_back)))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')
model.fit(trainX, trainY, epochs=10, batch_size=1, verbose=2)
```

```
Epoch 1/10
- 16s - loss: 0.0101
Epoch 2/10
- 15s - loss: 0.0027
Epoch 3/10
- 15s - loss: 0.0024
Epoch 4/10
- 15s - loss: 0.0023
Epoch 5/10
- 15s - loss: 0.0022
Epoch 6/10
- 15s - loss: 0.0021
Epoch 7/10
- 15s - loss: 0.0021
Epoch 8/10
- 15s - loss: 0.0020
Epoch 9/10
- 15s - loss: 0.0020
Epoch 10/10
- 15s - loss: 0.0020
```

```
Out[36]: <keras.callbacks.History at 0x7fa465b372e8>
```

Chapter 04: implementation and development

```
import pandas
from sklearn import model_selection
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn import preprocessing

# dataset
array = dataframe.values

X = array[:,0:1]
Y = array[:,1]

# normalize the dataset
le = preprocessing.LabelEncoder()
Y=le.fit_transform(Y)

np.random.seed(7)
# prepare models
models = []
models.append(('LR', LogisticRegression()))
models.append(('KNN', KNeighborsClassifier()))
models.append(('CART', DecisionTreeClassifier()))

# evaluation of each model
scoring = 'neg_mean_squared_error'
for name, model in models:
    kfold = model_selection.KFold(n_splits=5)
    cv_results = model_selection.cross_val_score(model, X, Y, cv=kfold, scoring=scoring )
    msg = "%s: %f" % (name, math.sqrt(cv_results.mean()*(-1)))
    print(msg)
    print("-----")
```

***** Australia Wind and Solar *****

LR: 227.809936

KNN: 178.026160

CART: 158.819308

**** RMSE results of MLs of Wind onshore power ****

***** Australia *****

LR: 86.033131

KNN: 66.738699

CART: 66.682462

**** RMSE results of MLs of Solar power ****

***** Australia *****

Chapter 04: implementation and development

***** Netherlands Wind and Solar *****

LR: 252.606339

KNN: 167.130502

CART: 152.107203

*** RMSE results of MLs of Wind onshore power ***
Netherlands

LR: 128.891194

KNN: 113.472356

CART: 107.757327

*** RMSE results of MLs of Solar power ***
Netherlands

4Conclusion

Python programming environment has given us many advantages in order to get the best results, which we have achieved through repeated experiments.

It should be noted that We might get better results than we did, this is thanks to the ability of LSTM to improve by playing on many parameters such as increasing the size of the window like look_back argument ,also we can reshape the data, like permutation between the time steps and the features or the add of other hidden layers but we have not tried them yet.

General Conclusion

General Conclusion

Bibliography

Bibliography

Bibliography

- [1] Workbook, Historical Data. "BP Renewable energy on Review " 2017.
- [2] Potter CW, Archambault A, Westrick K. Building a smarter smart grid through better renewable energy information. In Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES 2009 Mar 15 (pp. 1-5). IEEE.
- [3] Cecati C, Citro C, Siano P. Combined operations of renewable energy systems and responsive demand in a smart grid. IEEE transactions on sustainable energy. 2011 Oct;2(4):468-76.
- [4] Mathiesen BV, Lund H, Connolly D, Østergaard PA, Möller B. The design of Smart Energy Systems for 100% renewable energy and transport solutions. In Proceedings of the 8th Dubrovnik Conference for Sustainable Development of Energy, Water and Environment Systems. Dubrovnik, Croatia 2013 Sep (pp. 22-27).
- [5] Komor P, Molnar T. background paper on distributed renewable energy generation and integration.
- [6] Zomchak L, Umrysh H. modeling and forecasting of meat and eggs produsing in Ukraine with seasonal arima-model. Agricultural and Resource Economics International Scientific E-Journal. 2017 Sep 20;3(3):16-27.
- [7] Shafiullah GM, Oo AM, Ali AS, Wolfs P. Smart grid for a sustainable future. Smart Grid and Renewable Energy. 2013;4(1):23-34.
- [8] Cunjiang Y, Huaxun Z, Lei Z. Architecture design for smart grid. Energy Procedia. 2012 Jan 1;17:1524-8.
- [9] An Introduction to the Renewable Energy Resources- Shiva Gorjian – Tarbiat Modares University - June 2017.
- [10] U.S. Department of Energy, 20% Wind Energy by 2030, p. 63, July 2008.
- [11] Machine Learning Techniques for Supporting Renewable Energy Generation and Integration-Kasun S. Perera - Institute Center for Smart and Sustainable Systems.
- [12] Deep Learning –Ian Goodfellow, Yoshua Bengio, and Aaron Courville: The MIT Press, 2016, 800 pp, ISBN: 0262035618
- [13] Deep Learning for solar power forecasting — An approach using Auto Encoder and LSTM Neural Networks - Universität Kassel - André Gensler.
- [14] S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd ed. Upper Saddle River,NJ, USA: Prentice Hall PTR, 1998.
- [15] Szegedy C., Ioffe S., Vanhouche V., and AlemiA. Inception-v4, inception-resnet and the impact of residual connections on learning. Arxiv, 1602.07261, 2016.
- [16] Internet of things Role in the renewable energy resources- Al-Ali – 2016.
- [17] IoT in Smart Cities - Rolland Vida – 2016 .
- [18] Introduction to Big Data - Xiaomeng Su.
- [19] An Introduction to Big Data - An Introduction to Big Data – 2018.
- [20] Srivastava, S., & Lesmann, S. (2018). A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. Solar Energy, 162, 232-247.
- [21] Liang, S., Nguyen, L., & Jin, F. (2018, December). A Multi-variable Stacked Long-Short Term Memory Network for Wind Speed Forecasting. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 4561-4564). IEEE.
- [22] Gensler, A., Henze, J., Sick, B., & Raabe, N. (2016, October). Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks. In 2016 IEEE international conference on systems, man, and cybernetics (SMC) (pp. 002858-002865). IEEE.

Bibliography

- [23] Gensler, A., Henze, J., Sick, B., & Raabe, N. (2016, October). Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks. In 2016 IEEE international conference on systems, man, and cybernetics (SMC) (pp. 002858-002865). IEEE.
- [24] Xue, G., Pan, Y., Lin, T., Song, J., Qi, C., & Wang, Z. (2019). District Heating Load Prediction Algorithm Based on Feature Fusion LSTM Model. *Energies*, 12(11), 2122.
- [25] Eze, E. C., & Chatwin, C. R. (2019). Enhanced recurrent neural network for short-term wind farm power output prediction. *IJRDO-Journal of Applied Science*, 5(2), 28-35.
- [26] Hourly Day-Ahead Wind Power Prediction Using the Hybrid Model of Variational Model Decomposition and Long Short-Term Memory.
- [27] Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. (2018). Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies*, 11(7), 1636
- [28] "" Adele Kuzmiakova " Short-term Memory Solar Energy Forecasting at University of Illinois
- [29] Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*.
- [30] Sak H, Senior A, Beaufays F. Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. arXiv preprint arXiv:1402.1128. 2014 Feb 5.
- [31] Graves A. Generating sequences with recurrent neural networks. arXiv preprint arXiv:1308.0850. 2013 Aug 4.
- [32] Olah C. Understanding lstm networks. GITHUB blog, posted on August.2015 Aug 27;27:2015.
- [33] White T. Hadoop: The definitive guide. " O'Reilly Media, Inc."; 2012 May 19.
- [34] Zaharia M, Chowdhury M, Franklin MJ, Shenker S, Stoica I. Spark: Cluster computing with working sets. *HotCloud*. 2010 Jun 22;10(10-10):95.
- [35] Data Platform : Open Power System Data, https://data.open-power-system-data.org/time_series/.
- [36] Data Platform Open Power System Data https://data.open-power-system-data.org/weather_data/2018-09-04.
- [37] Du KL, Swamy MN. Fundamentals of machine learning. In *Neural Networks and Statistical Learning 2014* (pp. 15-65). Springer, London.
- [38] Han J, Pei J, Kamber M. *Data mining: concepts and techniques*. Elsevier;2011 Jun 9. Li H, Liu Z. Performance prediction and optimization of solar water heater via a knowledge based machine learning method. In *Handbook of Research on Power and Energy System Optimization 2018* (pp. 55-74). IGI Global.
- [39] Shearer C. The CRISP-DM: The new blueprint for Data Mining. *Journal of Data Warehousing*. 2000;5(4).
- [40] : python , retrieved from site : <https://python.com>
- [41] : python , retrieved from site : <https://www.tensorflow.org/install/>
- [42] : about Jupyter Notebook retrieved from site : <https://jupyter.org/>
- [43] : about Pandas retrieved from site <https://fr.wikipedia.org/wiki/Pandas>.
- [44] : about NumPy retrieved from site <https://en.wikipedia.org/wiki/NumPy>.

Bibliography

- [45] : about Matplotlib retrieved from site <https://fr.wikipedia.org/wiki/Matplotlib>
- [46] : about Matplotlib retrieved from site <https://matplotlib.org/>
- [47] : about Keras retrieved from site <https://pypi.org/project/Keras/>
- [48] : about Keras retrieved from site <https://en.wikipedia.org/wiki/Keras>
- [49] about sklearn retrieved from site: <https://fr.wikipedia.org/wiki/Scikit-learn>
- [50] Pierre N : Outils pour le BigData 2019 .