Project Report

SUBMITTED BY:

RITUL KUMAR SHARMA (Team Leader)

SHUBHAM SINHA

ANKIT SHARMA

MANIK BATRA

PROJECT TITLE :

Predicting the energy output of wind turbine based on weather condition.

Project ID : *SPS_PRO_ 188*

Challenge Title :*IBM Hack Challenge 2020*

SUBMITTED TO:

TheSmartBridge

1. INTRODUCTION

1. **Overview**: We will be able to predict the energy output of wind turbine.

Services (Cloud function , Watson Studio). By the end of the project, we’ll learn best practices of combining Watson services, and how they can build interactive information retrieval systems with Cloud Functions + Watson Studio .

- **Project Requirements**: Python, IBM Cloud, IBMWatson
- **Functional Requirements**: IBM cloud
- **Technical Requirements**: ML,WATSON STUDIO,PYTHON
- **Software Requirements**: Watson assistant, PYTHON.
- **Project Deliverables**: Smartinternz
- **Project Team**: alphaMARS
- **Project Duration**:30 days

2. **Purpose:**

Wind speed/power has received increasing attention around the earth due to its renewable nature as well as environmental friendliness. With the global installed wind power capacity rapidly increasing, the wind industry is growing into a large-scale business. Reliable short-term wind speed forecasts play a practical and crucial role in wind energy conversion systems, such as the dynamic control of wind turbines and power system scheduling. A precise forecast needs to overcome problems of variable energy production caused by fluctuating weather conditions. Power generated by wind is highly dependent on the wind speed. Though it is highly non-linear, wind speed follows a certain pattern over a certain period of time. We exploit this time series pattern to gain useful information and use it for power prediction.

2. Literature Survey:

1. *Existing Problem:*

Wind power is getting more and more attention around the earth due to its renewable nature. With the global installed wind power capacity rapidly increasing, the wind industry is growing into a large-scale business. Reliable short-term wind speed forecasts play a practical and crucial role in wind energy conversion, such as the dynamic control of wind turbines and power system scheduling. A precise and accurate forecast needs to overcome problems of variable energy production caused by fluctuating weather conditions. Power generated by wind is highly dependent on the wind speed.Wind speed follows a certain pattern over a certain period of time.

2. Proposed solution:

We all know that it looks very simple but it is more complex and challenging than we think. Accurate and reliable wind speed forecasts are a significant challenge due to its high rates of change, highly nonlinear behavior with no typical patterns, and dependency on elevation, terrain, atmospheric pressure, and temperature, which results in large uncertainties of wind speeds. This makes it difficult for any machine learning model to figure out a pattern and give an accurate prediction. We made it easy to interpret this problem as time series forecasting problem because the wind follows a particular pattern for a certain period for like a day, month or year. Long Short-Term Memory (LSTM) machine learning model, which is best known for time series data prediction is used to learn these patterns in wind and make a prediction about power. This prediction problem was divided into two categories:

1. **Estimation**: Weather conditions like temperature, wind speed, pressure etc. determining the energy power prediction.
2. **Prediction**: Without knowing any details about the weather conditions predicting the power generation using the pattern which it has followed in a certain period of time.

3. Theoritical analysis

Long short-Term Memory Machine Learning Model

Long short-term memory (LSTM) units (or blocks) are a building unit for layers of a recurrent neural network (RNN). A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell is responsible for "remembering" values over arbitrary intervals of time; hence the word "memory" in LSTM. Each of the three gates can be thought of as a "conventional" artificial neuron, as in a multi- layer neural network: i.e., they compute an activation of a weighted sum.moreover, they can be thought as regulators of the flow of values that goes through the connections of the LSTM; hence the denotation "gate". There are connections between these gates and the cell.
An LSTM is well- suited to classify, process and predict time series given time lags of unknown size and duration between important events. LSTMs were developed to deal with the exploding and vanishing gradient problem when training traditional RNNs. Relative insensitivity to gap length gives an advantage to LSTM over alternative RNNs, hidden Markov models and other sequence learning methods in numerous applications.
This just a introduction about LSTM, for more architectural and mathematical details you can read from the links provided in references.

Note:

Look Back/Lag is a common term used in LSTM which signifies the number of steps (apart from the pattern learned) an LSTM model will use to predict the next result.
In all the plots which are shown below. X-axis represents the hours for which we are predicting and Y-axis represents the power generated by the system.

Blue: TruePowerGenerated **Orange: Predicted PowerGenerated**

Data for LSTM Experiments

Historical wind energy data is taken from NREL to do this analysis. 6 years of wind power generation data is used in this experiment. The data after pre-processing have details about timestamp, air temperature (C), pressure (atm), wind direction (deg), wind speed (m/s) and Power generated by the system (kW). We have hourly data for about 6years.

All the features are used for Estimation model and only time series features i.e. Date Time and power generated by the system are used for prediction experiments.

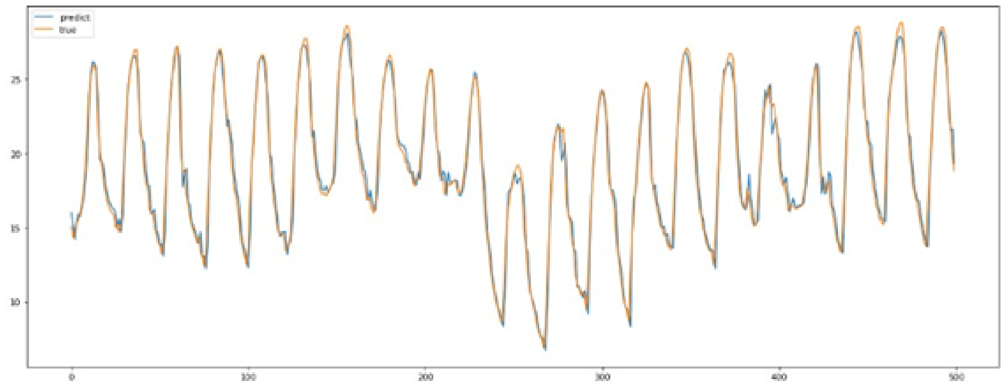
Estimation

Estimation is all about predicting wind power generation given the current wind direction. Current wind and temperature conditions are given, this makes this problem a bit easy for a model like LSTM which looks at the current state of the weather and the previous trend which the weather is following to predict the power generated by the system Estimations models are useful if we get the weather information about the present day or the future publicly using machine learning with certain accuracy. Then this model can be used to be the perfect estimation of power generated by the system.

4. Experimental investigation

Experiment 1:

Six years hourly data was divided into 70-30 train test batch for this experiment. That means 4 years of data was used to predict 2 years of wind power generation. Good result with root mean square error(RMSE) 1.242 and Variance 0.984 was observed for this experiment.




```
In [11]: print("Mean squared error: %.3f" % mean_squared_error(testY, yhat))
```

```
Mean squared error: 1.242
```

```
In [12]: print("Root mean squared error: %.3f" % sqrt(mean_squared_error(testY, yhat)))
```

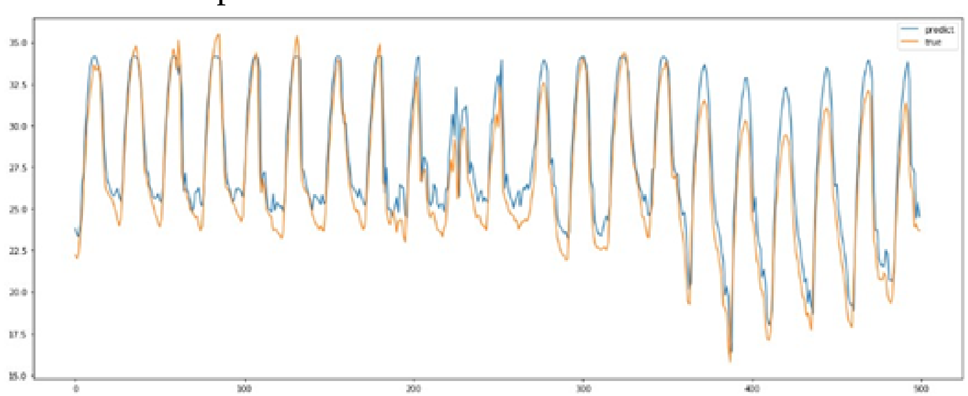
```
Root mean squared error: 1.115
```

```
In [13]: from sklearn.metrics import mean_squared_error, r2_score
         print('Variance : %.3f' % r2_score(testY, yhat))
```

```
Variance : 0.984
```

Experiment 2:

Six years hourly data was divided into 60-40 train test batch for this experiment. That means 3 years of data was used to predict 3 years of wind power generation. Good result with RMSE of 1.667 and Variance 0.969 was observed for this experiment.




```
In [11]: print("Mean squared error: %.3f" % mean_squared_error(testY, yhat))
```

```
Mean squared error: 2.778
```

```
In [12]: print("Root mean squared error: %.3f" % sqrt(mean_squared_error(testY, yhat)))
```

```
Root mean squared error: 1.667
```

```
In [13]: from sklearn.metrics import mean_squared_error, r2_score
         print('Variance : %.3f' % r2_score(testY, yhat))
```

```
Variance : 0.969
```

5. *Pediction*

Let's look into the pure time series analysis. In Prediction part, we are predicting the power generated by the system without any knowledge of the future weather. This is important because predicting the future weather is also a different prediction problem machine learning with its different set of challenges. We are not going to have any knowledge what wind speed is going to be or what temperature or pressure is going to be in the future. So, we try to predict the power only by analyzing pattern in the past data using LSTM. Data to this model will be Date time and Power generated by the system in the supervised form as required by the LSTM. LSTM will analyze the prior data and try to get useful knowledge about the patterns in previous data. And using that knowledge it is going to predict the results. Walk forward validation is used for predicting future values and evaluating results.

Data to LSTM for prediction:

	Power generated by system | (kW)
DateTime	
2007-01-01 00:00:00	33688.1
2007-01-01 01:00:00	37261.9
2007-01-01 02:00:00	30502.9
2007-01-01 03:00:00	28419.2
2007-01-01 04:00:00	27370.3

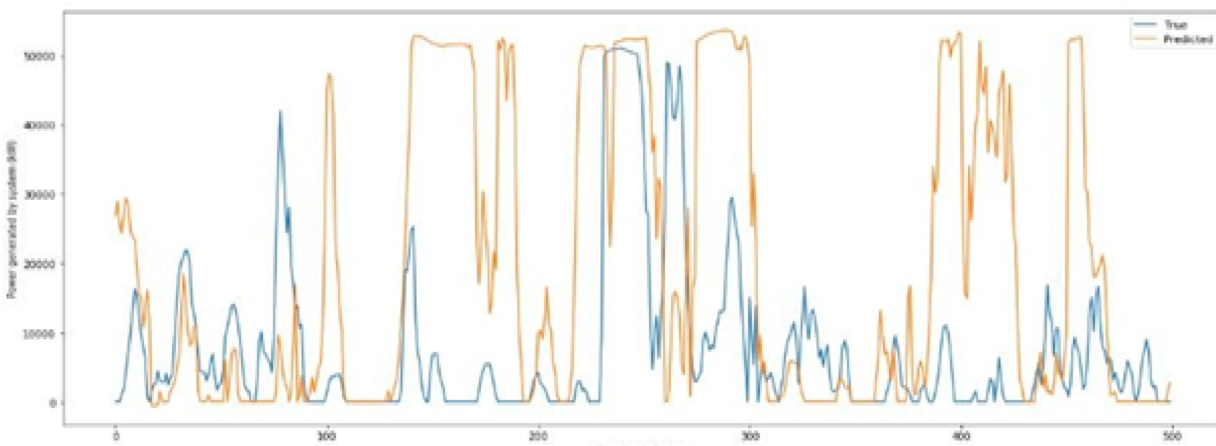
After forming the baseline LSTM model, we performed many different experiments to get the perfect look back and the neurons which are required by the LSTM. After getting the look back, neuron number and certain other parameters right for the model we performed a few experiments and predictions. The result of them as follows.

6. *Result:*

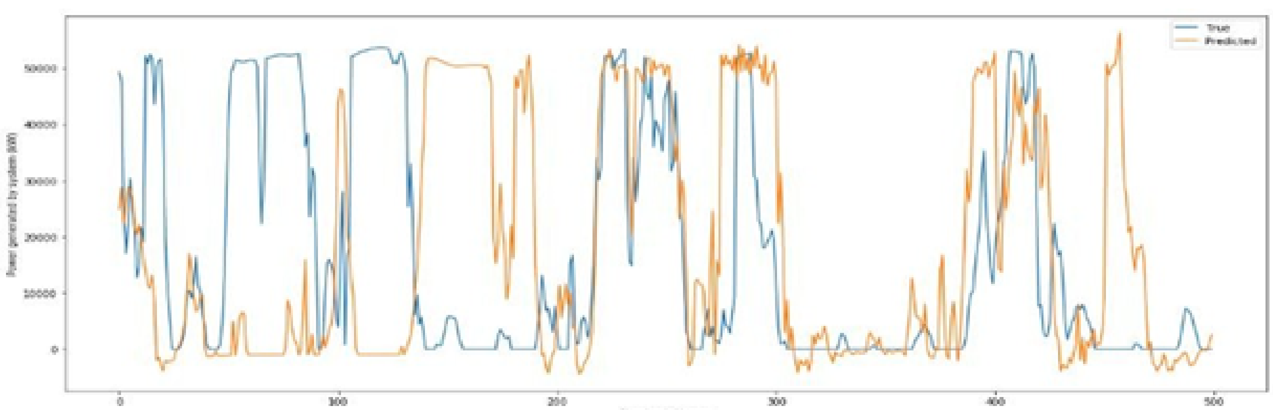
<u>Results of Experiments for Washington(WA) as follows:</u>

Experiment No.	Prediction duration	Result (Mean Absolute Percent Error)
1	24 hours	99.74
2	2 days	117.59
3	1 week	127.70
4	1 month	124.45

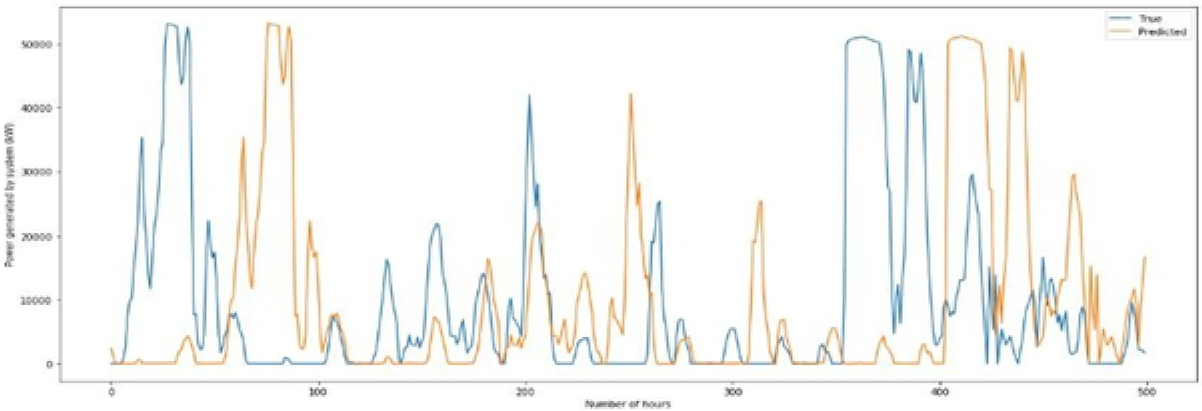
<u>Plots for Prediction:</u>



<u>1. Month ahead Prediction</u>



<u>2.Week aheadPrediction</u>



1. <u>3.Days aheadPrediction</u>

Out[10]:

DateTime	Air temperature | (°C)	Pressure | (atm)	Wind speed | (m/s)	Wind direction | (deg)	Power generated by system | (kW)
2007-01-01 00:00:00	10.926	0.979103	9.014	229	33688.1
2007-01-01 01:00:00	9.919	0.979566	9.428	232	37261.9
2007-01-01 02:00:00	8.567	0.979937	8.700	236	30502.9
2007-01-01 03:00:00	7.877	0.980053	8.481	247	28419.2
2007-01-01 04:00:00	7.259	0.979867	8.383	256	27370.3

7. *Advantages and disadvantages:*

<u>Advantages:-</u>

1. Once the wind turbine is built the energy it produces does not cause green house gases or other pollutants.

2.Wind turbines have a role to play in both the developed and third world.

3. Remote areas that are not connected to the power grid can use wind turbines to produce their own electricity

4.Many people find wind farms an interesting feature of the landscape.

<u>Disadvantages:-</u>

1. Wind turbines are noisy. Each one can generate the same level of noise as a family car travelling at 70 mph.

2. When wind turbines are being manufactured some pollution is produced. Therefore wind power does produce greenhouse gasses and pollution.

3. Many people feel that the countryside should be left untouched, without these large structures being built. The landscape should left in its natural form for everyone to enjoy.

4. Many people see large wind turbines as unsightly structures and not pleasant or interesting to look at. They disfigure the countryside and are generally ugly.

8. *Applicatons:*

As we all know Renewable energy is the future of energy and wind mills is one that would be mostly used as a renewable source of energy because it takes less space as compared to others, more efficient and doesn't harm the environment in any way, that's why this system will be high in demand for cost reduction in construction and maintenance when constructing it according to the whether of the given place and will also help in increasing the efficiency of the energy output.

9. *Conclusion*

In this study we showed that wind energy output can be predicted from publicly available weather data with accuracy at best 80% R2 on the training range and at best 85,5% on the unseen test data. We identified the smallest space of input variables (windGust2 and dewPoint), where reported accuracy can be achieved, and provided clear trade-offs of prediction accuracy for decreasing the input space to the windGust2 variable. We demonstrated that an off-the-shelf data modeling and variable selection tool can be used with mostly default settings to run the symbolic regression experiments as well as variable importance,variable contribution analysis, ensemble selection and validation.

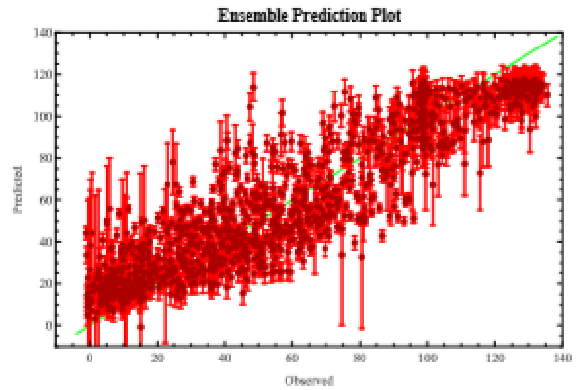


Figure 10: Ensemble prediction versus observed energy output in July (Test Data) of the final model ensemble. Whiskers correspond to ensemble disagreement measured as a standard deviation between predictions of individual ensemble members for any given input sample.

We are looking forward to discuss the results with domain experts and check the applicability of produced models in real-life for short term energy production prediction. We are glad that the presented framework is so simple that it can be used literally by everybody for predicting wind energy production on a smaller scale—for individual wind mills on private farms or urban buildings, or small wind farms. For future work, we are planning to study further the possibilities for longer-term wind energy forecasting.

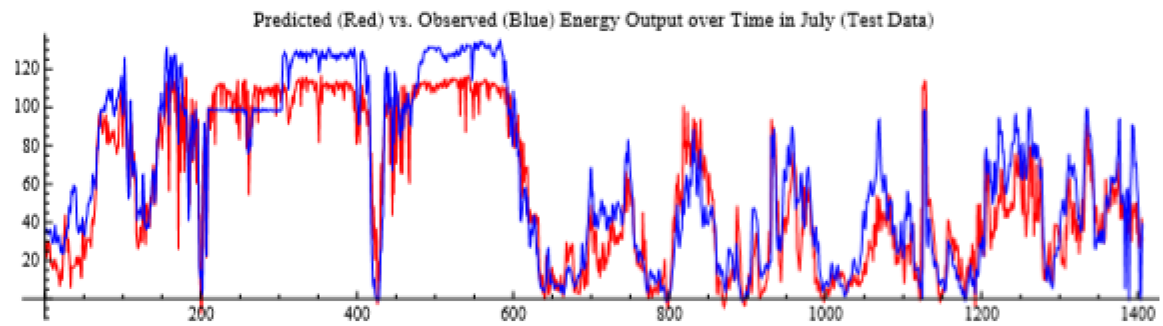


Figure 11: Ensemble Prediction versus Actual energy output over time on the Test Data.

10. *Future Scope:*

India is the home of 1.25 billion people i.e. 17.5% of the total world population, which makes it second most populous country in world. India has the second fastest growing economy of the world. India's substantial and sustained economic growth over the years is placing enormous demand on its energy resources. The electricity sector in India had an installed capacity of 253.389 GW as of August 2014 [1]. India became the world's third largest producer of electricity in the year 2013 with 4.8% global share in electricity generation surpassing Japan and Russia. Power development in India was first started in 1897 in Darjeeling, followed by commissioning of a hydro-power station at Sivasamudram in Karnataka during 1902. Thermal power stations which generate electricity more than 1000 MW are referred as Super Thermal Power Stations. India'selectricity generation capacity additions from 1950 to 1985 were very low when compared to developed nations. Since 1990, India has been one of the fastest growing markets for new electricity generation capacity [2]. India's electricity generation capacity has increased from 179 TW-h in 1985 to 1053 TW-h in 2012. Wind energy is indigenous and helps in reducing the dependency on fossil fuels. Wind occurrence is due to the differential heating of the earth's crust by the sun.

11. *Bibliography:*

APPENDIX

Source Code:

1. **Node Red(flow.json)**


```
[
  {
    "id": "7253a121.16642",
    "type": "tab",
     "label":
    "Flow1","disa
    bled":false,"in
    fo":""
   },
  {
    "id": "b1b11140.4e4ef",
    "type": "inject",
    "z": "7253a121.16642",
    "name": "",
    "topic": "",
```

"payload": "Hello

Node-RED!","payloadType"

: "str",

"repeat": "",

"crontab": "",

"once":

false,"once

Delay": "","x":

141,

"y": 61,

"wir

es":

[[

"2371449b.4bf2cc"

]

]
},

{

"id": "f2f2649a.0d0d98",

"type": "debug",

"z": "7253a121.16642",

"name": "",

"active":

true,"to side

bar":

true,"console

": false, "to

status": false,

"complete":"paylo

ad","target Type":

"msg","x":670,

"y":140,

"wires": []

},

{

"id": "e150cc25.d7aa",

"type": "function",

"z": "7253a121.16642",

"name": "input parsing",

"func": "msg.payload=msg.payload.text;\nreturn

msg;","outputs": 1,

"noerr": 0,

"x":270,

"y":240,

"wir

es":

[[

"2371449b.4bf2cc",

"49234419.64d55c"

]

]

},

{

"id": "d3215cca.6b835",

"type": "ui_form",

"z": "7253a121.16642",

"name": "",

"label": "",

"group": "528b62f3.5c783c",
"order": 6,

"width": "0",

"height":

"0","options": [

{

"label": "Enter your query",

"value": "text",

"type":

"text",

"required":

true, "rows":

null

}

],

"formValue": {

"text": ""

},

"payload": "",

"submit": "submit",

"cancel": "cancel",

"topic":"",

"x":107.5,
"y": 302,

"wir

es":

[[

"e150cc25.d7aa"

]

]

},

{

"id": "49234419.64d55c",

"type": "ui_text",

"z": "7253a121.16642",

"group":

"528b62f3.5c783c","ord

er": 1,

"width": "6",

"height": "1",

"name": "",

"label": "You",

"format":

"{{msg.payload}}","layout":

"col-center",

"x": 453.5,
"y": 306,

"wires": []

},

{

"id": "d2a61bef.00f978",

"type": "ui_text",

"z": "7253a121.16642",

"group":

"528b62f3.5c783c","ord

er": 2,

"width": "6",

"height": "1",

"name": "",

"label": "Bot",

"format":

"{{msg.payload}}","layout":

"col-center",

"x": 688.5,

"y": 237,

"wires": []

},

{
"id": "29aac8ee.b98c68",

"type": "function",

"z": "7253a121.16642",

"name": "parsing",

"func": "msg.payload = msg.payload.output.generic[0].text;\nreturn

msg;","outputs": 1,

"noerr": 0,

"x":470,

"y":160,

"wir

es":

[[

"f2f2649a.0d0d98",

"d2a61bef.00f978"

]

]

},

{

"id":

"2371449b.4bf2cc","type":

"watson-assistant-v2","z":

"7253a121.16642",

"name":"Myfirstassistant",

"service-endpoint":
"https://api.eu-gb.assistant.watson.cloud.ibm.com/instances/ae44476e-77ab-4578-9d11-23464ea66634",

"assistant_id":
"d3fa58b9-7946-4784-b120-5e025c58c9ba","debug":
false,
"restart": false,
"return_context":
true,"alternate_intents":
false,"multisession":
true, "timeout": "",
"optout-learning": false,
"x": 391.5,
"y": 100,

"wir
es":
[[

"29aac8ee.b98c68","c
af5ac15.c978d"
]

]

},
{

"id": "caf5ac15.c978d",

"type": "debug",

"z": "7253a121.16642",

"name": "",

"active":

true,"tosideba

r":

true,"console

": false,

"tostatus":

false,

"complete":

"payload","targetType":

"msg",

"x": 557.5,

"y": 61,

"wires": []

},

{

"id": "528b62f3.5c783c",

"type": "ui_group",

"z": "",

"name": "",
"tab":

"27e055be.28bb7a","order"

: 1,

"disp": true,

"width":

"6","collapse"

: false

},

{

"id": "27e055be.28bb7a",

"type": "ui_tab",

"z": "",

"name": "Product",

"icon":
"dashboard","order": 2, "disabled": false, "hidden": false
}
]

2. You can go through from .ipynb files for Machine Learning Source Code.