**ANALYZE SMS MESSAGES WITH WATSON KNOWLEDGE STUDIO**

**1. INTRODUTION:**

**1.1 OVERVIEW**

Text analysis is a [machine learning](https://monkeylearn.com/machine-learning/) technique that allows companies to automatically understand text data, such as tweets, emails, support tickets, product reviews, and survey responses. You can us text analysis to extract specific information, like keywords, names, or company information from thousands of emails, or [categorize survey responses](https://monkeylearn.com/blog/sentiment-analysis-of-survey-responses/) by sentiment and topic.

Text analysis delivers [qualitative results](https://monkeylearn.com/blog/qualitative-feedback/) and text analytics delivers quantitative results. If a machine performs text analysis, it identifies important information within the text itself, but if it performs text analytics, it reveals patterns across thousands of texts, resulting in graphs, reports, tables etc. [Text analysis tools](https://monkeylearn.com/blog/text-analysis-tools) allow businesses to structure vast quantities of information, like emails, chats, social media, support tickets, documents, and so on, in seconds rather than days, so you can redirect extra resources to more important business tasks.

[Text classification](https://monkeylearn.com/text-classification/) is the process of assigning predefined tags or categories to unstructured text. It's considered one of the most useful natural language processing techniques because it's so versatile and can organize, structure, and categorize pretty much any form of text to deliver meaningful data and solve problems. [Natural language processing](https://monkeylearn.com/natural-language-processing/) (NLP) is a machine learning technique that allows computers to break down and understand text much as a human would.

**1.2 PURPOSE**

Businesses are inundated with information and customer comments can appear anywhere on the web these days, but it can be difficult to keep an eye on it all. Text analysis is a game-changer when it comes to detecting urgent matters, wherever they may appear, 24/7 and in real time. By training text analysis models to detect expressions and sentiments that imply negativity or urgency, businesses can automatically flag tweets, reviews, videos, tickets, and the like, and take action sooner rather than later.

**1.3 PROBLEM STATEMENT**

The SMS messages in this code pattern are related to merchants offering special offers to their customers. With NLU, you can extract some general information from each text, but you might want to add the capability to extract additional specific data, such as what the offer is, who the merchant is, how long the offer is valid, and what the merchant’s phone number and website is. You can accomplish this by loading sample messages into WKS and training it to recognize entities within each text. The result is a model that you can then use to process additional messages.

For example:

PIZZA! Don't Cook Wednesdays are here! Get 50% off a Medium Pizza.

Offer available for single Pizza in-store and two for Home Delivery.

Walk-In/Call @ 555-555-5555

The example above has a few interesting entities which could not be extracted with conventional NLP techniques, but by using Watson services we can find out the following:

1. What is the offer?
2. Who is the merchant?
3. What is the offer name?
4. What is the offer's validity period?
5. What is the merchant's phone number?
6. What is the merchant's website?

**2. LITERATURE SURVEY**

**2.1 PROJECT OBJECTIVE**

After completing this code pattern, the user will learn how to:

* Upload a corpus with WKS
* Import types to WKS
* Use WKS to create a model
* Deploy a WKS model to NLU
* Call NLU APIs with a WKS model specified

**2.2 EXISTING PROBLEM**

For example, if we have 1000 of application for an interview or insurance. We have to provide the documents. The organization has to analyze the documents based their specific requirements and their requirements. It is difficult for humans to analyze all the application. It may takes lots of days to complete the task. Existing system is manual analysis of the human of the documents, messages, feedback, questions & answers, ..etc.

**2.3 PROPOSED SYSTEM**

We are providing a machine which we analyze the documents, messages, feedbacks, tweets, question & answers,..etc. If the user provide the inputs, the message will be analyze based on the type of the message, the message that represents and confidence.

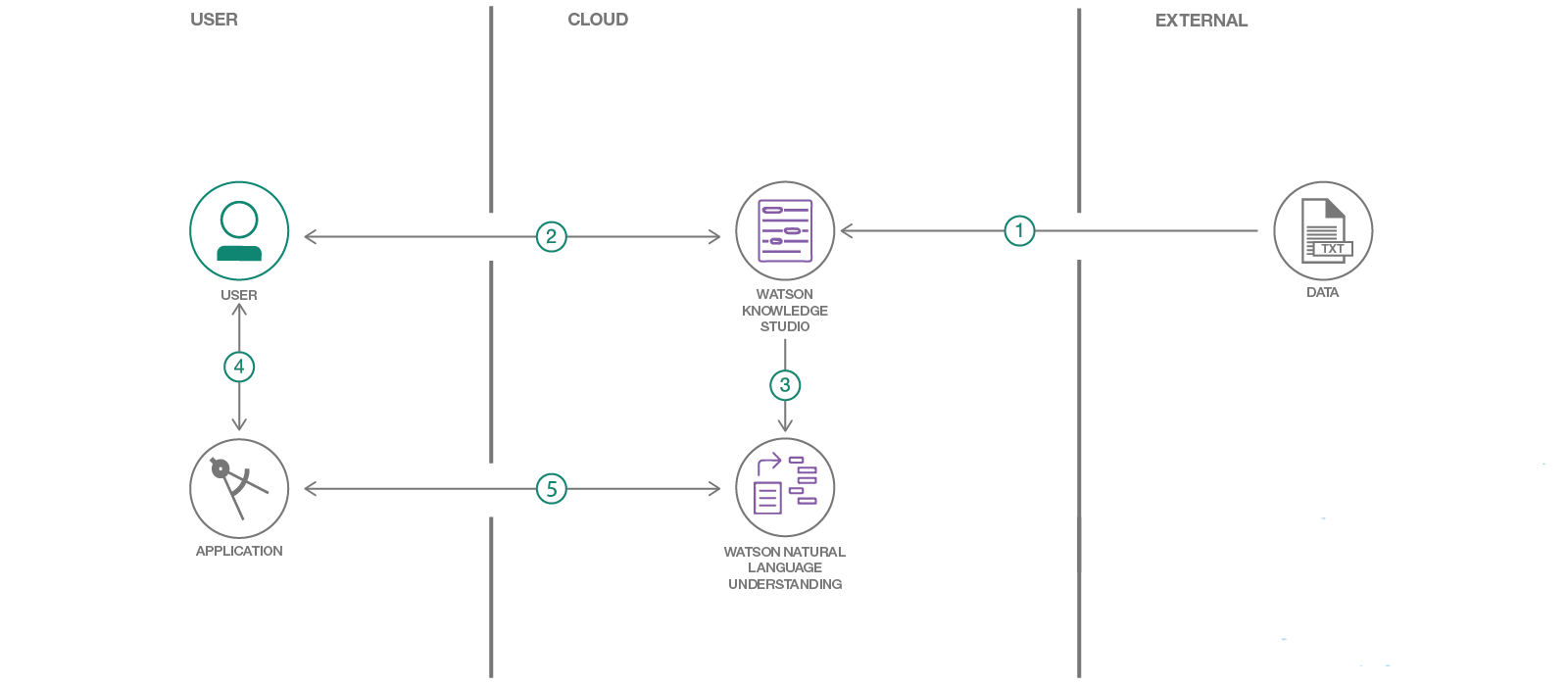
**2.4 PROJECT FLOW**

The project flow contains the following steps

* Create IBM account
* Log in to IBM Account
* Create Watson Knowledge Studio and Watson Natural language Understanding Services
* Load entity type system and Training data files into Watson Knowledge Studio.
* Generate model by training and evaluating data.
* The WKS model is deployed to Watson NLU.
* Build Node-RED application (UI)
* A user provides an SMS message to the app for analysis.
* The SMS message is analyzed by Watson NLS for processing and returns extracted domain-specific entities

**3. THEORITICAL ANALYSIS**

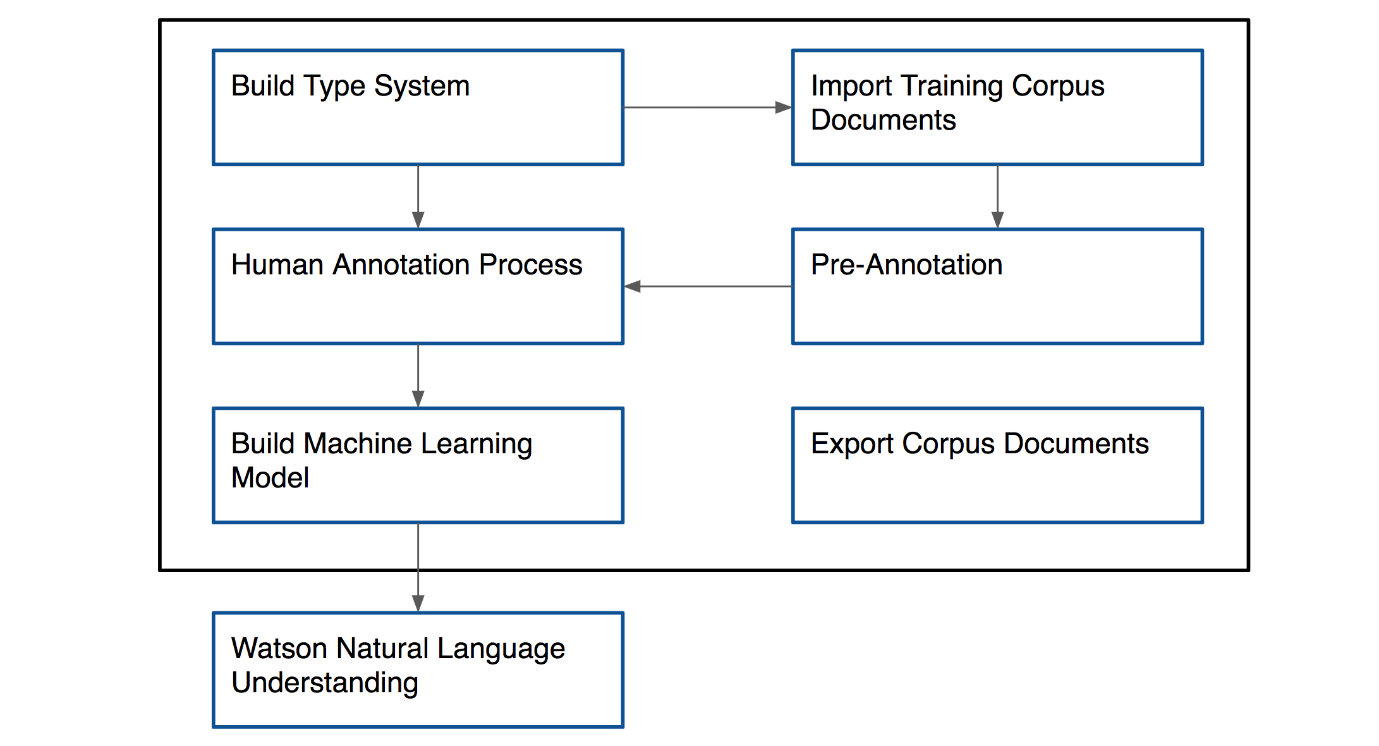
**3.1 TECHNICAL ARCHITECTURE:**



1. Load type system and corpus files into Watson Knowledge Studio.
2. A user generates a model by training and evaluating data.
3. The WKS model is deployed to Watson NLU.
4. A user provides an SMS message to the app for analysis.
5. The SMS message is analyzed by Watson NLS for processing and returns extracted domain specific entities based on the WKS model are returned.

**3.2 WORKING OF WATSON KNOWLEDGE STUDIO**

In short, a type system is built and supporting documents are uploaded that have domain specific wording. From here a model must be built to properly understand the documents, this is where the annotations come in. Once the corpus and annotations are set you are free to create a model and deploy it to a Watson Natural Language Understanding instance.



**3.3 HARDWARE/SOFTWARE DESIGNING**

**Software Requirements:**

* Os- windows XP 7,8,10
* IBM account

**Prerequisites :**

For this project, you should create the following services

* **Watson Knowledge studio:-** Teach Watson the language of your domain with custom models that identify entities and relationships unique to your industry, in unstructured text. Use the models in Watson Discovery, Watson Natural Language Understanding, and Watson Explorer.
* **Watson Natural Language Understanding:-** An IBM Cloud service that can analyze text to extract meta-data from content such as concepts, entities, keywords, categories, sentiment, emotion, relations, semantic roles, using natural language understanding.
* **Node-RED :-** Node-RED is a flow-based development tool for visual programming developed originally by IBM for wiring together hardware devices, APIs and online services as part of the Internet of Things. Node-RED provides a web browser -based flow editor, which can be used to create JavaScript functions.

**4. IMPLEMENTATION**

In this activity, you will be building a custom model for domain-specific entities.  
In this milestone you should complete the following tasks:

1. Create a workspace
2. Upload entity types
3. Upload training data
4. Annotate the documents
5. Train and evaluate the model
6. Deploy model in NLU

### 1.*CREATE A WORKSPACE*

### Step1:- A workspace defines all the resources that are required to create a machine learning model, including training documents, the type system, dictionaries, and annotations that are added by human annotators.

From Knowledge, studio service page click on Launch Watson Kanowledge Studio

### Step2:-

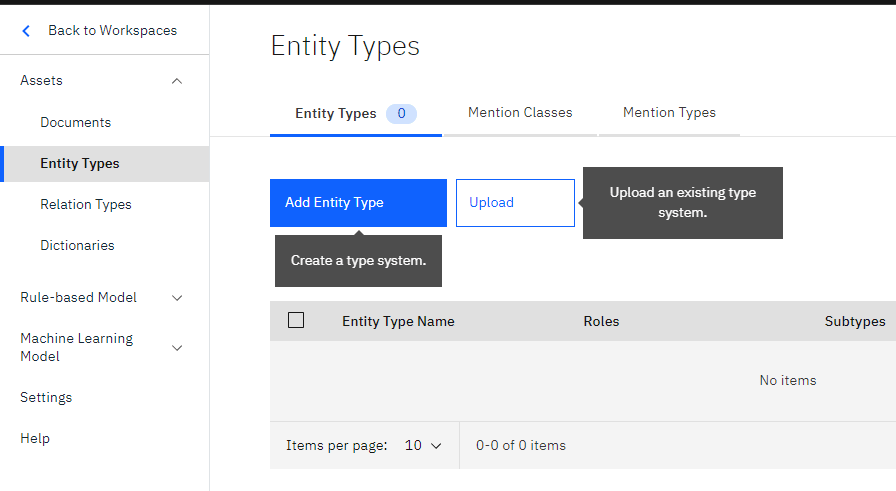
### Click on create entities and relation types

### Step3:-

### Give a name for ex: Sms-analysis and click on create

**Step4:-**

You will be redirected to a page which looks  like below

****

### 2 *UPLOAD ENTITY TYPES*

### Step1:-

For this project, we have created a JSON file named Entitytypes file which has all the entities that can be extracted from an SMS for example (offer, merchant, coupon code, etc.,)

Select the Entity types option from the left  pane. Click on upload

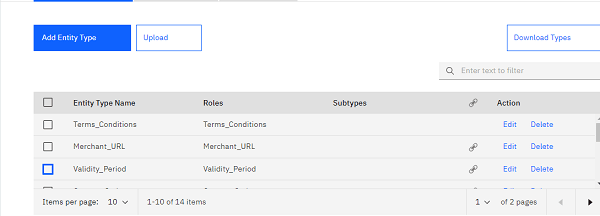
**Step2:-**

Pop-up displays, browse for **entytypes.json** file and click on upload

**Step3:-**

you can see all the entities are uploaded to type system

After you finish uploading entity types to the type system, you can begin adding documents(training data) to your workspace.



### 3. *UPLOAD TRAINING DATA*

**Step1:-**

Now click on Documents from the left pane  you can see there is no training data

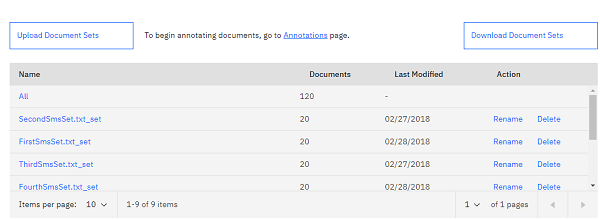
Now click on upload document sets

**Step2:-**

Browse for training data.zip file and click on upload

**Step3:-**

You can see all the training data is been uploaded



### 4. *CREATE ANNOTATION TASK*

**Step1:-**

Now let us annotate the documents (train data)  with the uploaded entities

click on the annotation page

**Step2:-**

Select Automation tasks and select Add Task

**Step3:-**

Now Give a name to the task as shown in the figure and click on create annotation sets

**Step4:-**

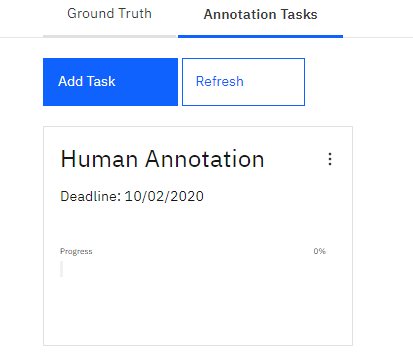
A window pops up select All as an option  in Base set, select the owner and give a name as **NewAnnotation**s to your  annotations

Click on generate

**Step5:-**

Now click on **save**at the top right corner,  you will redirect to a page which looks like this

You can see progress is zero  , now lets annotate the document sets



### 5.*ANNOTATE DOCUMENT SETS*

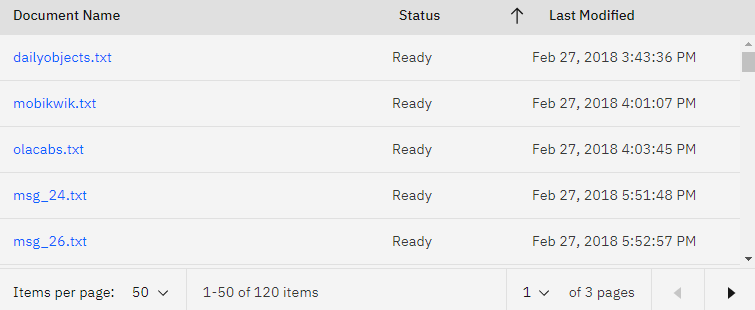
### Step1:-

Double click on the Human automation task

You will redirect to a page where click on the annotate option

### Step2:-

You will redirect to a page which looks like this

****

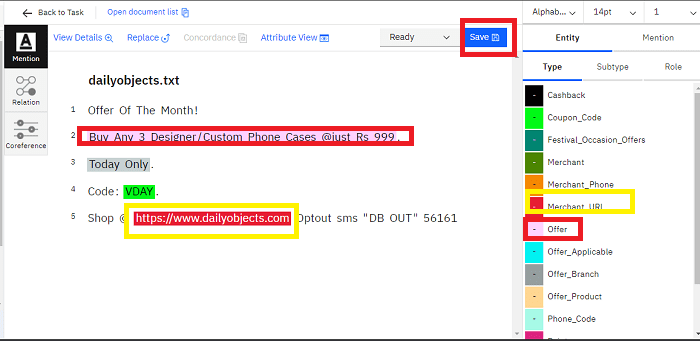
### Step3:-

You can see the status as ready for all the documents, the reason is IBM has an internal annotator which will make all the annotations ready for us. But if you want to create your own annotations click on any of the documents for example click on daily object.txt

you will get redirected to a page where you will find some text for annotation and entity types lies towards the right

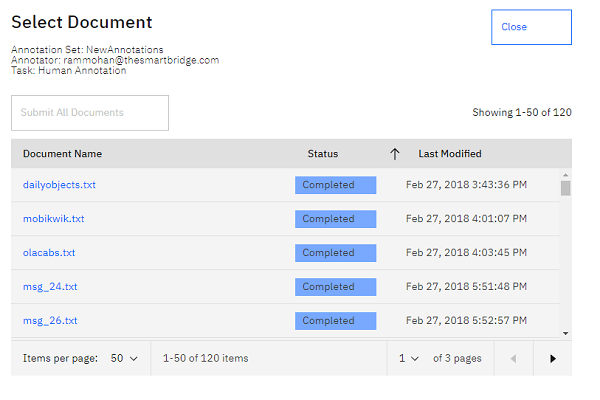
Select the text and give the appropriate annotation to the selected text , once annotation is done click on save repeat this for all the documents.

once annotation is done click on  Open Doccumentlist. you will get redirected to the previous page

****

### Step4:-

Click on submit All documents and  you will get redirected to this page, click on close

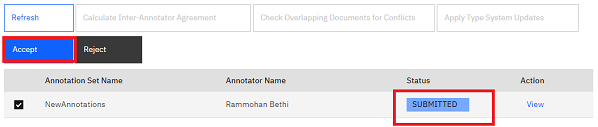
****

### Step5:-

Now click on the back to task option from top.You will get redirect to a page looks below.

Select the checkbox and click on accept.The submitted status changes to completed

Now you have completed your Annotations lets train these annotations

****

### 6.*TRAIN THE ANNOTAATIONS*

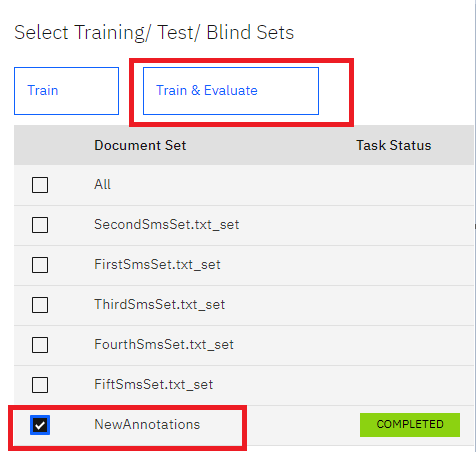
### Step1:-

Click on Performances then you will find an option called to train and evaluate .

Click on train and Train and evaluate

### Step2:-

select the annotation set that you have submitted and click on the train and evaluate button. it will take about 10 to 15 minutes for training

****

### 7.*DEPLOY THE IN NLU SERVICE*

### Step1:-

Click on machine Learning --> select version --> click on create version

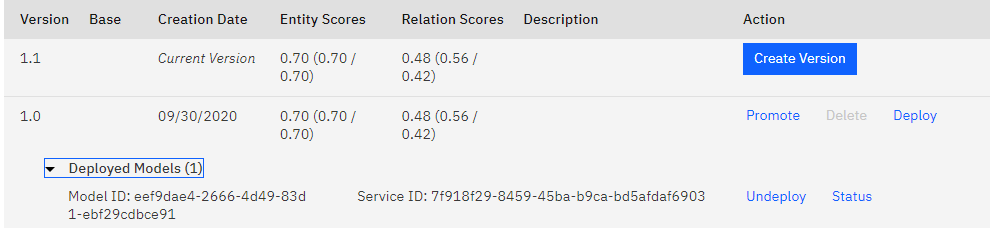
a window pops up, click on ok

Version is created. Click on deploy

**Step2:-**

once the model is deployed a model ID is generated

we can use this model id in we or android applications

  
**BUILD APPLICATION**

**Step -1 :** Open the Node-Red and click the Go to your Node-RED flow editor.

**Step-2 :** Search the flow and drag the form button. Create a new group which is named as the SMS and edit it. Create a variable to store the value given by the user. Then click done.

**Step-3:** Select the function button and drag it to the flow. Connect the form and function. Create a variable with msg.url. Here the url is from the natural understating language and model id is from the version which was deployed. Name the function as NLU URL. After updating Click done

**Step-4:** Select the http.request and drag it to connect the NLU URL. Give the method name as GET. Enable the authentication and give the user name as api key password is taken from the Natural understanding language key.

**Step-5:** Select the function and drag it to connect the http.request. Give a msg.payload variable and assign msg.payload.entities. Click done.

**Step-6:** Select the template and drag it to connect the function. Select template type and widget and write the code for the application. Click done.

**Step-7:** Select the debug and drag it to connect the NLU URL. Select the debug option and output as url.

**Step-8:** Select inject and drag it. Give the name as time stamp and give 0.1 second

**Step-9:** Select the function node and drag it to connect the timestamp and the http.request nodes. Give the url similar to the NLU URL. Then click done.

**Step-10:** Select the debug node and connect it to http.request node after dragging it. Update it as in the figure and click done.

**Step-11:** Deploy the model and select the dashboard then click pop up button it will directly to the output.

**Step-12:** We have to provide the input. Then the output will be displayed.

**5. RESULTS**

This Analyze SMS message with Watson Knowledge Studio contributes of three steps. They are:

* Input message is taken from the user which sms message like offers, discounts.
* The data will be trained by the knowledge studio.
* Output of classification of the text is given to the user along with the confidence.

**6. ADVANTAGES AND DISADVANTAGES**

**Advantages**

### Helps identify the root of a problem (or source of satisfaction).

### Enables emerging trends to surface that many feedback surveys limit or restrict.

### Issues can be prioritised quickly and efficiently.

### Customers’ ideas and suggestions materialise, leading to an enhanced digital experience.

### Accurate and easy usable interaction.

### User friendly application.

**Disadvantages**

* There may loss of any authorized or secured data.
* can be difficult to automate or computerize.
* The emotions will not be understanded clearly to the text analysis.

**7. APPLICATIONS**

### Knowledge Management

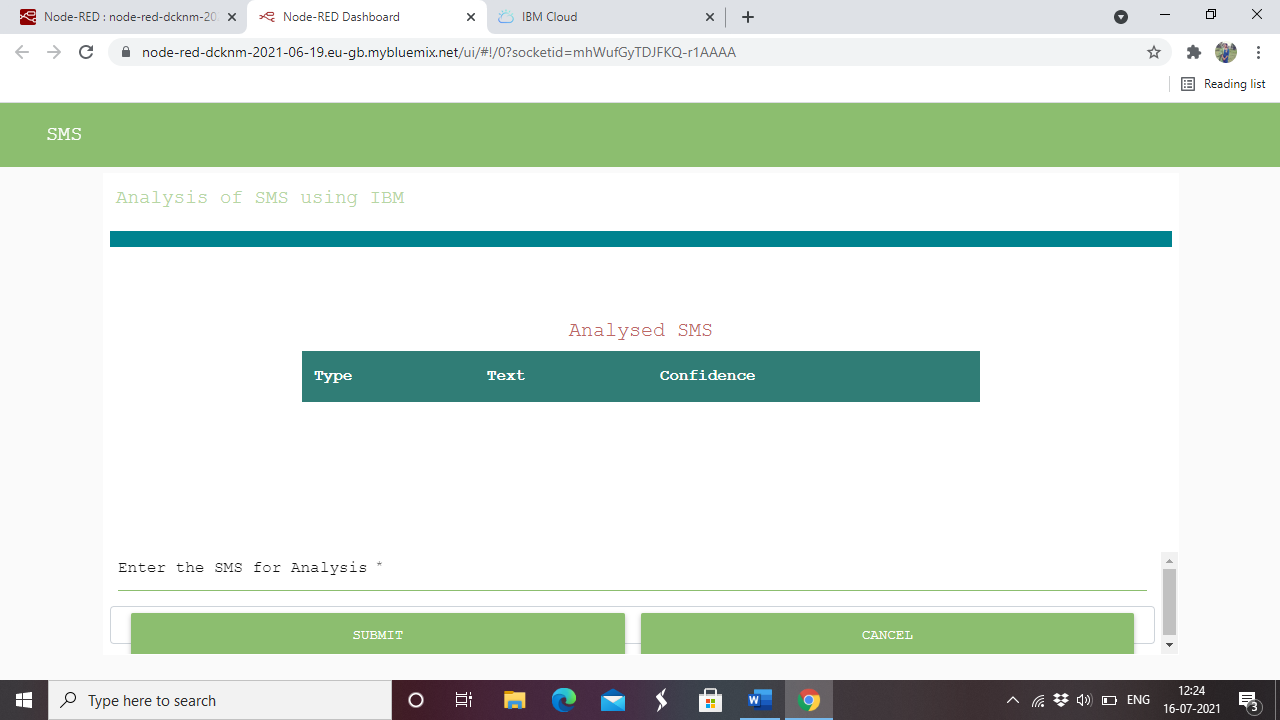
### Social Media Analysis

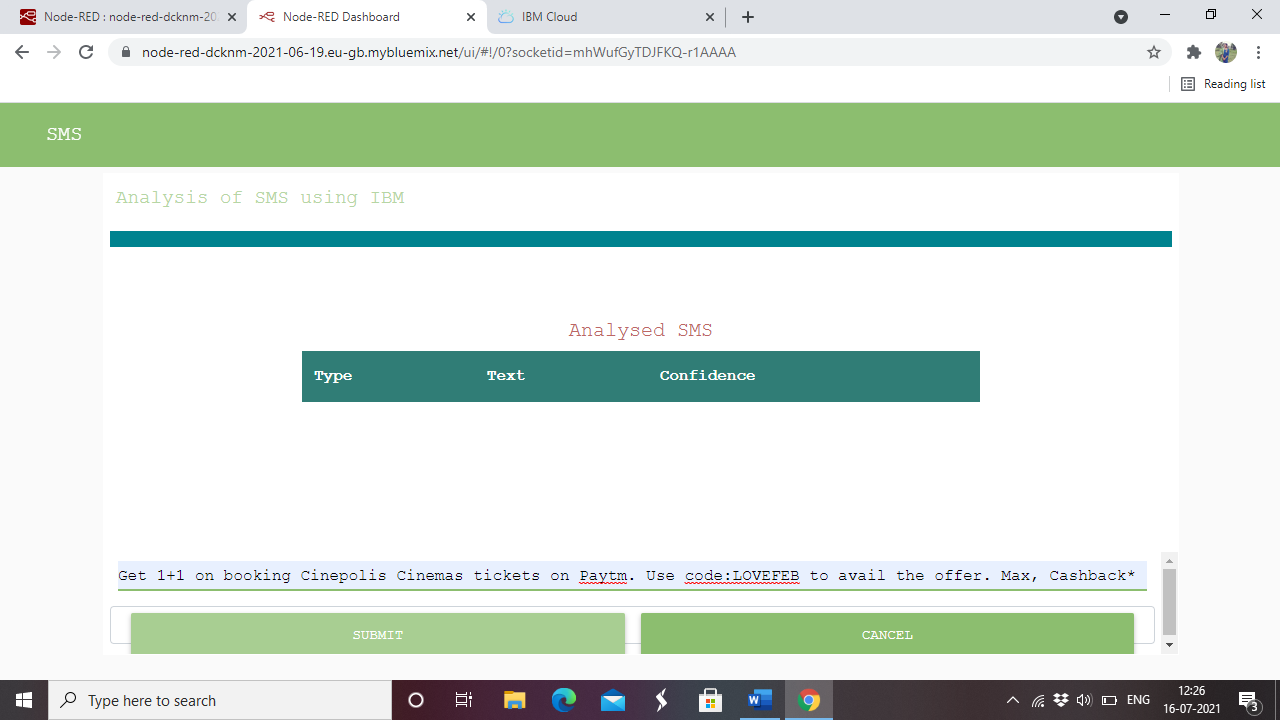
### Customer Care Service

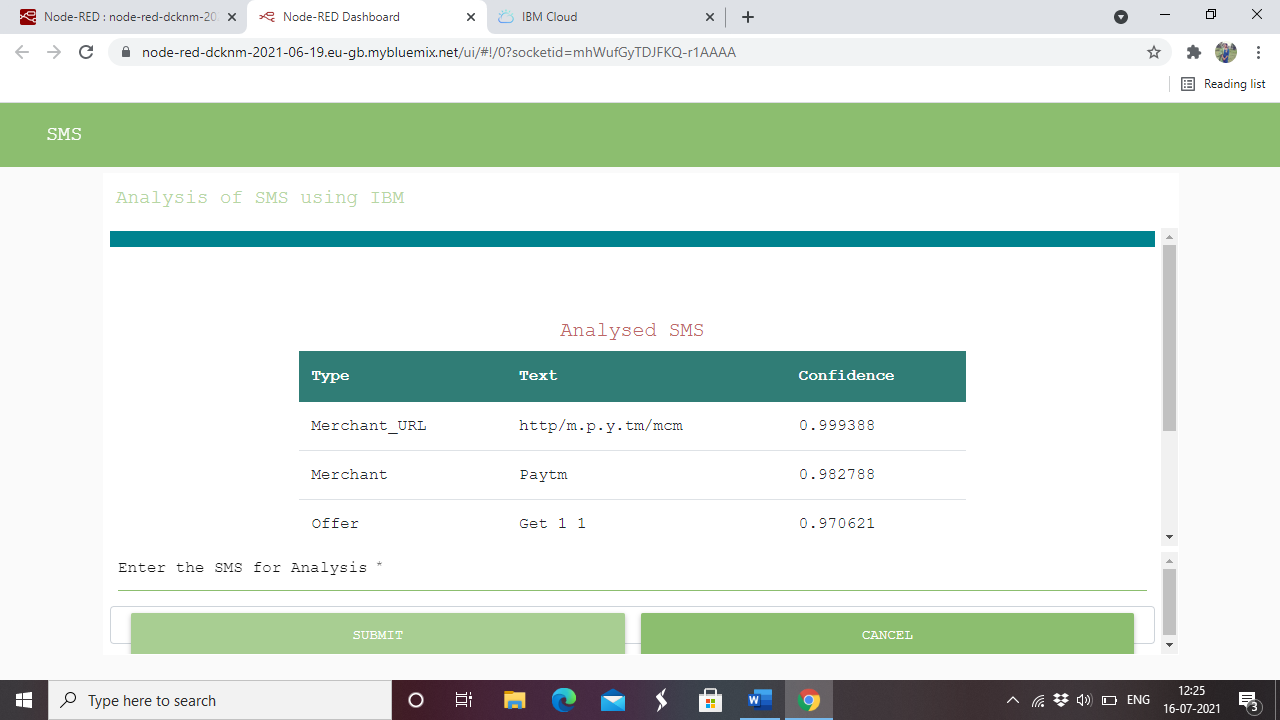
### Fraud Detection

## Marketing analysis

**8. SCREEN SHOTS**







**9.CONCLUSION**

In today’s competitive world, it is difficult for the everyone to spare the time on analysing the data. This application will made everyone easy to analyse the texts. ML text analysis is a technology that is used in various applications. Special models help to teach the machine to work with such data and draw valuable conclusions from it. All in all, it can be a valuable technique for generating insights for your product or for your business.

**10. FUTURE SCOPE**

* Can be done as mobile application
* Can be done using flask
* Developing the application more interactively for analysing documents, tweets, question & answers,…. Etc.

**11. BIBILOGRAPHY**

**REFERENCES**

## Build a custom model to better categorize SMS message content using Watson Knowledge Studio and Watson Natural Language Understanding:- <https://developer.ibm.com/technologies/artificial-intelligence/patterns/analyze-sms-messages-with-watson-knowledge-studio/>

# Datasets for Analyzing SMS messages with Watson Knowledge Studio:- <https://github.com/IBM/sms-analysis-with-wks>

1. [Watson Knowledge Studio - Overview](https://www.ibm.com/cloud/watson-knowledge-studio" \t "_blank):- <https://www.ibm.com/cloud/watson-knowledge-studio>
2. [Watson Natural Language Understanding](https://www.ibm.com/cloud/watson-natural-language-understanding):- <https://www.ibm.com/cloud/watson-natural-language-understanding>
3. [Build a recommendation engine with Watson Natural Language](https://developer.ibm.com/technologies/artificial-intelligence/tutorials/build-a-recommendation-engine-with-watson-natural-language-understanding) Understanding:- <https://developer.ibm.com/technologies/artificial-intelligence/tutorials/build-a-recommendation-engine-with-watson-natural-language-understanding>