SMS analysis using Watson Knowledge Studio

Introduction

SMS analysis using the Watson Knowledge Studio project examines the textual messages and extracts meta-data like keywords, concepts, categories, emotions, semantic rules, sentiments, and any domain-specific requirements. The custom model uses a machine learning technique called Text Analysis. The flexible and collaborative environment provided by the Watson Knowledge studio makes it possible to train the custom model based on automatic annotations and easily apply it to any new document.

Manually processing and organizing text data takes time, it's tedious, inaccurate, and it can be expensive if you need to hire extra staff to sort through text. Since the custom machine learning model is trained on real-life text samples, it makes it easier to collect and extract the entity types from a large textual document, thus reducing cost and essentially time.

Theoritical Analysis

Natural Language Processing

Natural language processing (NLP) refers to the branch of computer science, and more specifically, the branch of Artificial Intelligence concerned with giving computers the ability to understand the text and spoken words in much the same way human beings can.
NLP combines computational linguistics rule-based modeling of human language with statistical, machine learning, and deep learning models. Together, these technologies enable computers to process human language in the form of text or voice data and to ‘understand’ its full meaning, complete with the speaker or writer’s intent and sentiment.

NLP drives computer programs that translate text from one language to another,

respond to spoken commands, and summarize large volumes of text rapidly even in real-time. There's a good chance you've interacted with NLP in the form of voice-operated GPS systems, digital assistants, speech-to-text dictation software, customer service chatbots, and other consumer conveniences. But NLP also plays a growing role in enterprise solutions that help streamline business operations, increase employee productivity, and simplify mission-critical business processes.

Natural Language Understanding

Natural Language Understanding is a subtopic of Natural Language Processing in artificial intelligence that deals with machine reading comprehension. The IBM Watson Natural Language Understanding service enables developers to extract insights from unstructured text to power a new generation of cognitive applications. It is used to analyze text and extract meta-data like keywords, concepts, sentiments, relations, keywords, categories, semantic rules, or any domain-specific customizations

Watson Knowledge Studio

Its is an interface where we can teach the IBM Watson the language of our domain with custom models that identify entities and relationships unique to our industry.
Its is used to build models in a collaborative environment designed for both developers and domain experts, without needing to write code.
We can create machine learning models that understand linguistic meaning or a rule based model that finds entities and models based on your domain
We can train our models to identify certain entites and define custom relation types.

Node-RED

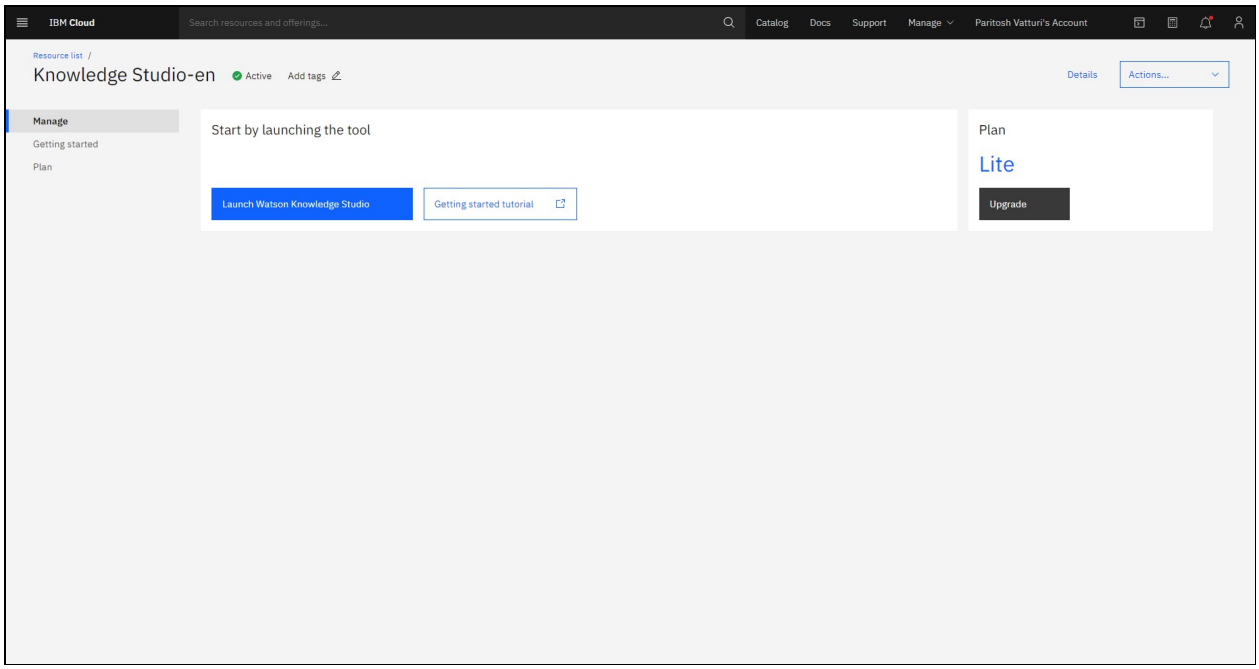
Node-RED is a programming tool for wiring together hardware devices, APIs, and online services in new and interesting ways. It provides a browser-based editor that makes it easy to wire together flows using the wide range of nodes in the palette that can be deployed to its runtime in a single click.JavaScript functions can be created within the editor using a rich text editor. A built-in library allows you to save useful functions, templates, or flows for re-use. The flows created in Node-RED are stored using JSON which can be easily imported and exported for sharing with others.

Procedure

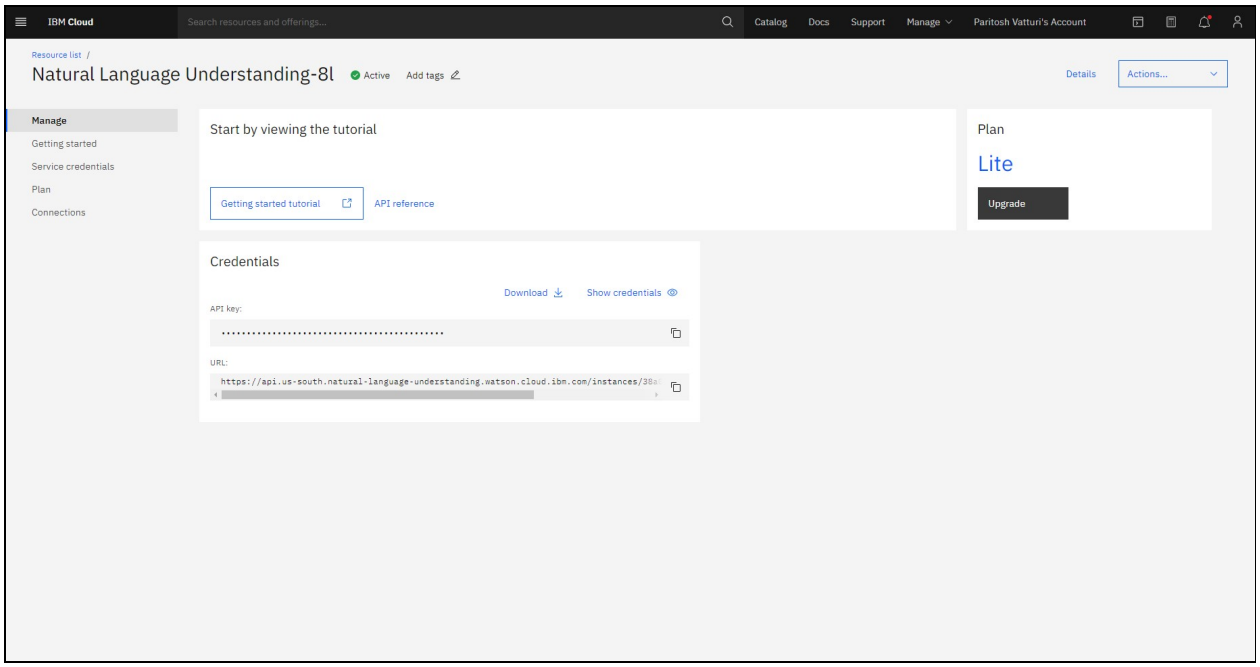
The project flow contains the following steps -

- Create IBM account

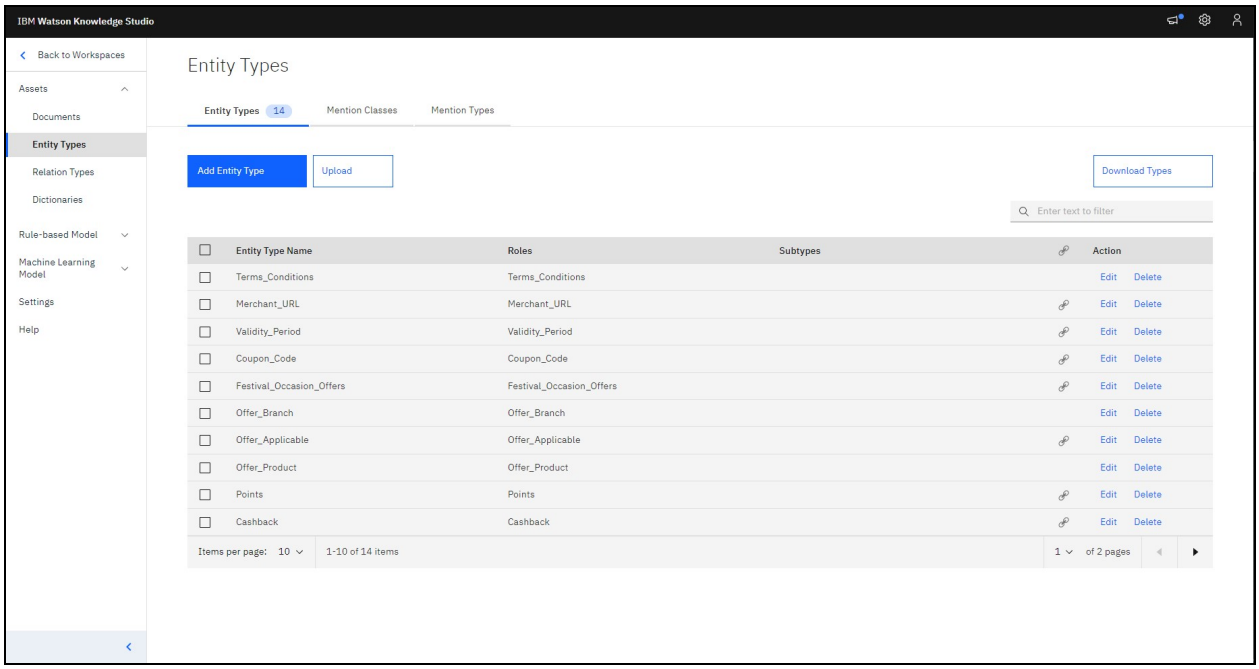
- Create Watson Knowledge Studio Service



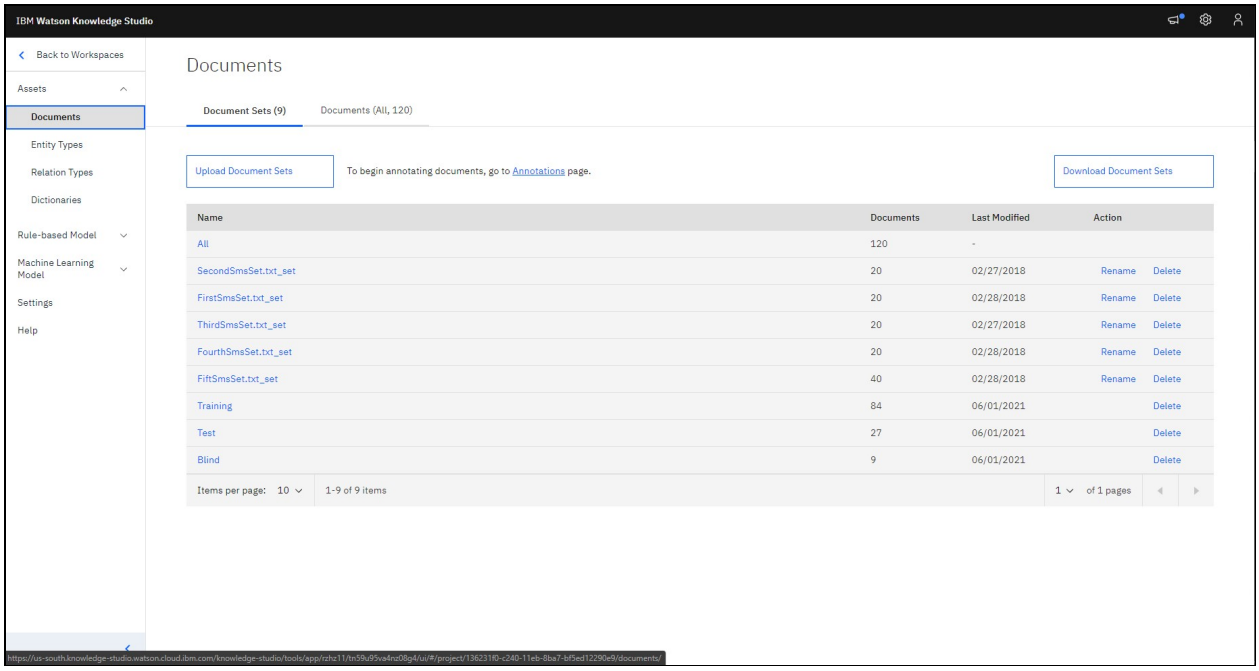
Create Watson Natural language Understanding Service



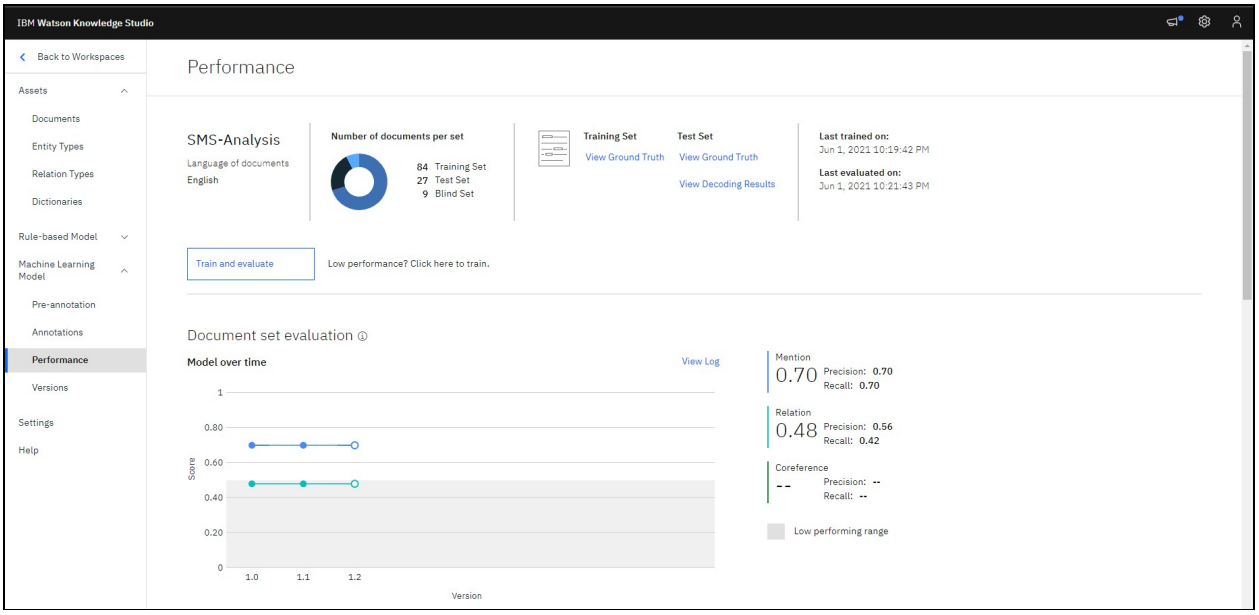
- Load entity type system in Watson Knowledge Studio



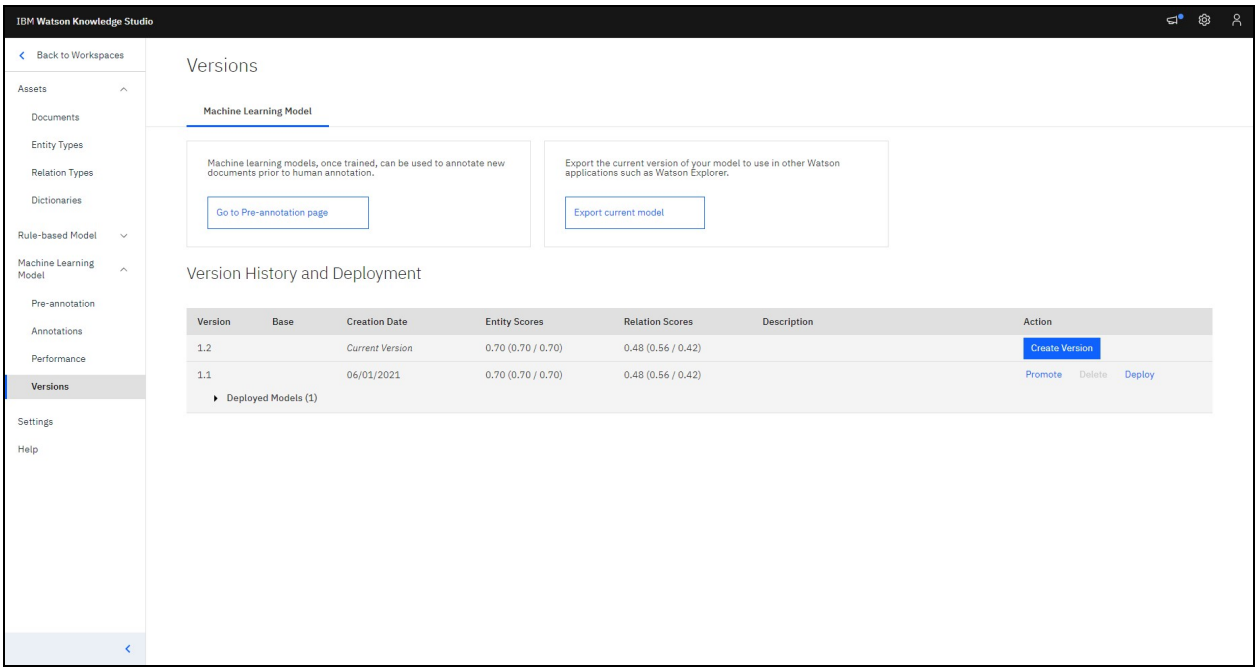
- Load training data files into Watson Knowledge Studio



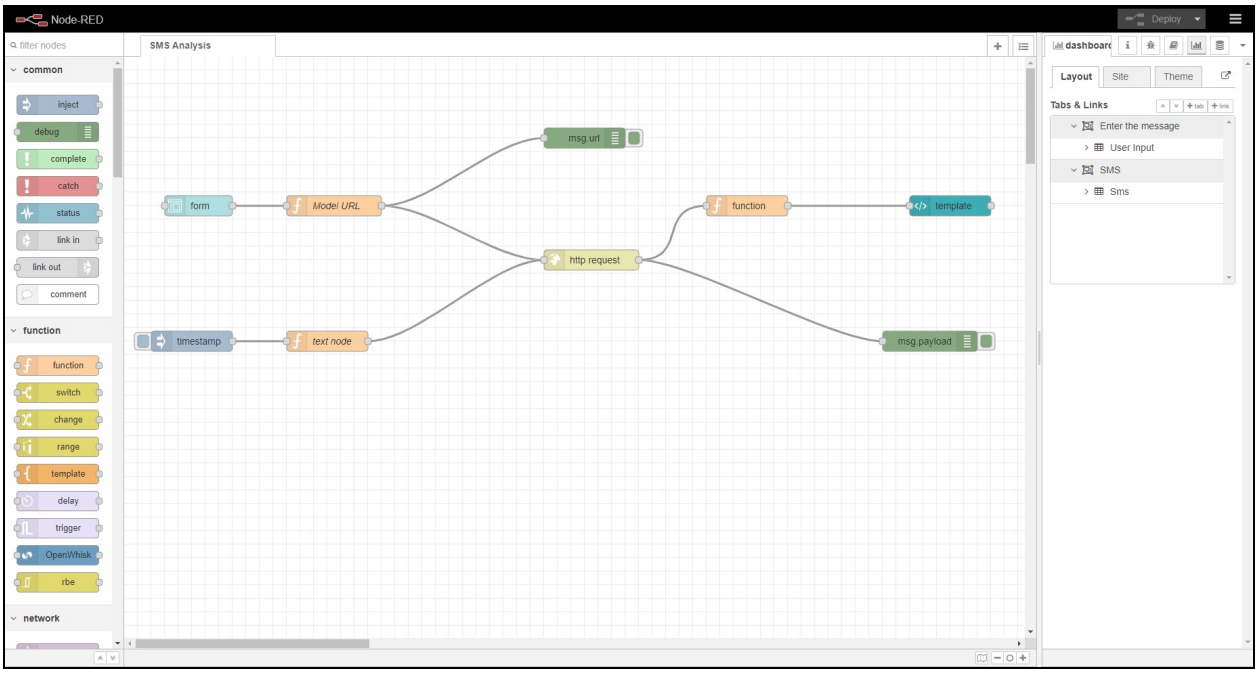
- Generate model by training and evaluating data.



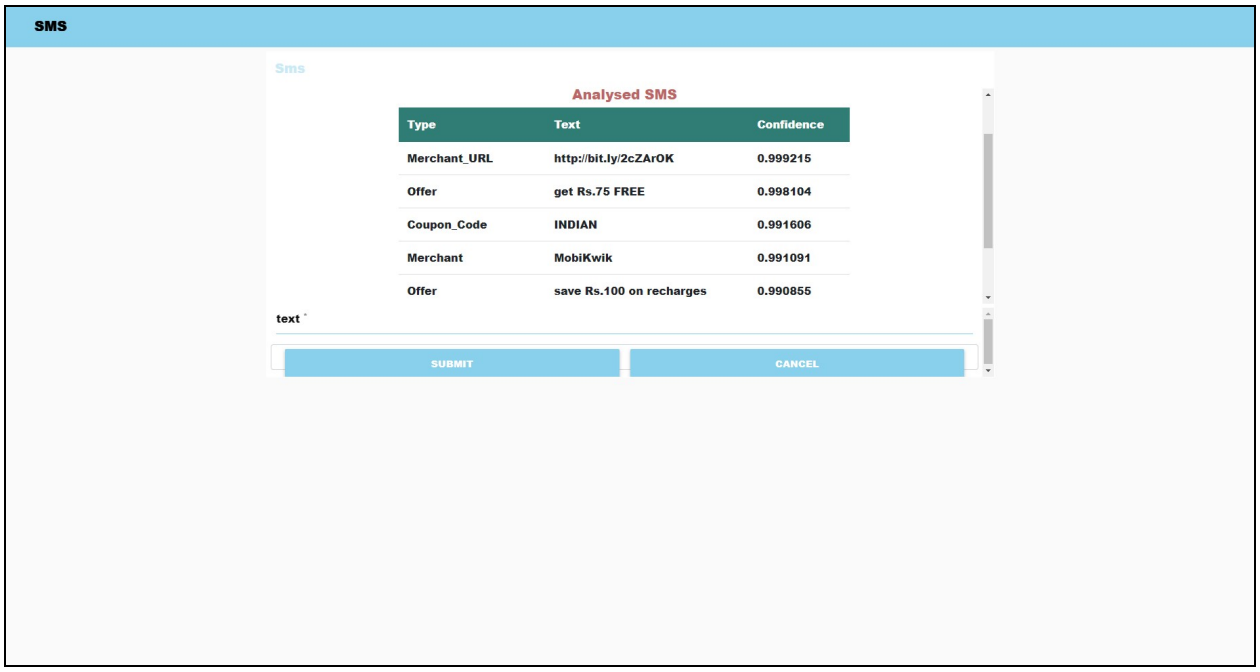
- The Watson Knowledge Studio model is deployed to Watson NLU.



- Build Node-RED application (UI)



- The SMS message is analyzed by Watson NLS for processing and returns extracted domain-specific entities



Conclusion

Our team has successfully completed the project SMS analysis using Watson Knowledge Studio. By all counts and proven results, this model is efficient in extracting entity types from text messages.
We would like to thank all the mentors from SmartBridge for their support and guidance throughout the course of the project.

Thank you