Q. Take a data set of your choice and analyse it or perform some operations on it in MonkeyLearn tool.

What is MonkeyLearn?

MonkeyLearn is a **Machine Learning** platform for **Text Analysis**. It allows our users to easily get actionable data from raw text. For example, you can detect topic or sentiment expressed in texts like tweets, chats, reviews, articles, and more.

MonkeyLearn provides:

* A Graphical User **Interface that allows users to easily create and test customized machine learning models** to solve particular problems.
* **Publicly available and pre-trained models** for common problems (sentiment analysis, topic detection, etc).
* A scalable cloud computing platform where machine learning algorithms can be trained and ran instantly without installing or deploying any software.
* An [API](https://monkeylearn.com/api/) and SDKs ([Python](https://github.com/monkeylearn/monkeylearn-python), [Ruby](https://github.com/monkeylearn/monkeylearn-ruby), [Node](https://github.com/monkeylearn/monkeylearn-node), [Java](https://github.com/monkeylearn/monkeylearn-java), and [PHP](https://github.com/monkeylearn/monkeylearn-php)) that allows users to integrate the MonkeyLearn cloud computing engine with any software project, using any programming language.
* [Documentation](http://help.monkeylearn.com/) and [blog](https://monkeylearn.com/blog/) to provide additional content around guides and use cases.

One of the stand-out features in MonkeyLearn is that you can train a highly dependable Machine Learning model on the fly with your particular data. More accuracy is gained by using texts from your own domain and building a model with your specific criteria in mind.

Models in MonkeyLearn are organized into two families:

* **Classification:** models that take text and return labels or categories.
* **Extraction**: models that extract particular data within a text.

So in this assignment i am performing sentimental analysis of tripadvisor\_hotel\_reviews.csv data set. This data set has 20491 reviews of different hotel ,so we perform sentimental analysis that either person feel good about hotel or not.

Now, more than ever, it’s key for companies to pay close attention to the [voice of customer (VoC)](https://monkeylearn.com/voice-of-customer/) to improve their products.

Product managers need insights that will help them develop a robust product roadmap; it’s about providing customers with what they actually want, rather than with what businesses think they need.

A good place to start collecting [product feedback](https://monkeylearn.com/blog/product-feedback/) is from online review sites (such as Capterra, G2Crowd, and Google Play). But manually analyzing this [unstructured data](https://monkeylearn.com/unstructured-data/) would take far too long.

That’s where [sentiment analysis](https://monkeylearn.com/sentiment-analysis/) can help to:

* Understand what your customers like and dislike about your product.
* Compare your product reviews with those of your competitors.
* Get the latest product insights in real-time, 24/7.
* Save hundreds of hours of manual data processing.

Sentiment analysis is the automated process of understanding the sentiment or opinion of a given text. You can use it to automatically analyze product reviews and sort them by *Positive, Neutral, Negative*.

The best part. You can start analyzing your product reviews for sentiment right away with [MonkeyLearn](https://monkeylearn.com/" \t "_blank), a no-code platform that’s simple and quick to use.

Follow our guide, below, to learn how to run sentiment analysis on your product reviews.

1. [Gather product reviews](https://monkeylearn.com/#gather)
2. [Run a sentiment analysis on product reviews](https://monkeylearn.com/#sentiment-analysis)
3. [Visualize the results of your sentiment analysis](https://monkeylearn.com/#visualize)

1. Gather Product Reviews

This section provides a high-level explanation of how you can automatically gather your product reviews.

Product reviews are everywhere on the Internet. You might stumble upon your brand’s name on Capterra, G2Crowd, Siftery, Yelp, Amazon, and Google Play, just to name a few, so collecting data manually is probably out of the question.

Thankfully, the bleak days of copying and pasting are long gone. [Web scraping](https://en.wikipedia.org/wiki/Web_scraping) can help to automate and streamline this whole process. Web scrapers are used to collect information from across the Internet. These tools simulate how people surf the web to gather specific data from different websites. In essence, they automatically find what you would otherwise have to copy and paste manually from any given website.

Generally speaking, web scraping tools can be grouped into two distinct categories:

So i gathered hotel reviews from kaggle .You can also make or download data set.

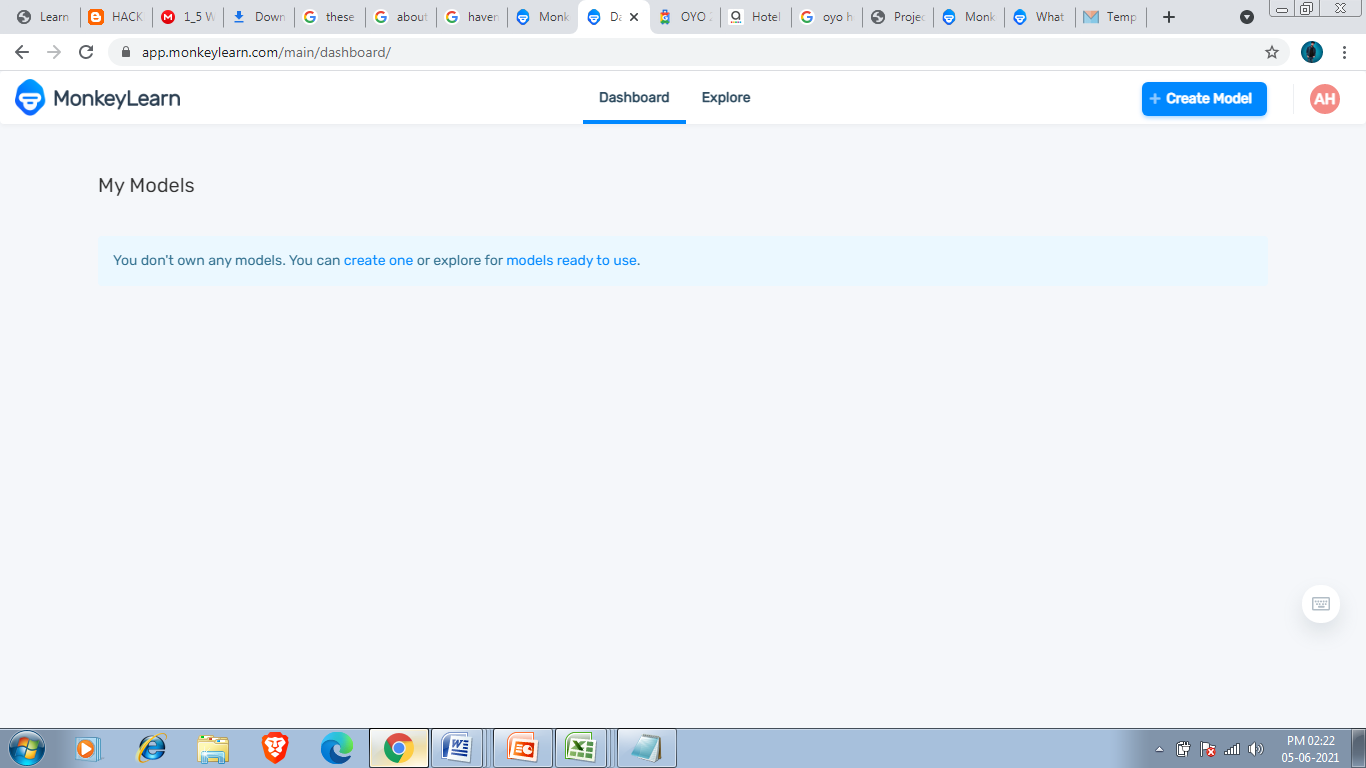
So its sentimental analysis help me to understand that according to customers which hotel is best and which hotel is worst.It will help in my business growth.

**Create a Sentiment Analysis Classifier**

Here you’ll learn how to create and test a sentiment analysis model for analyzing product reviews in six easy steps. Check it out:

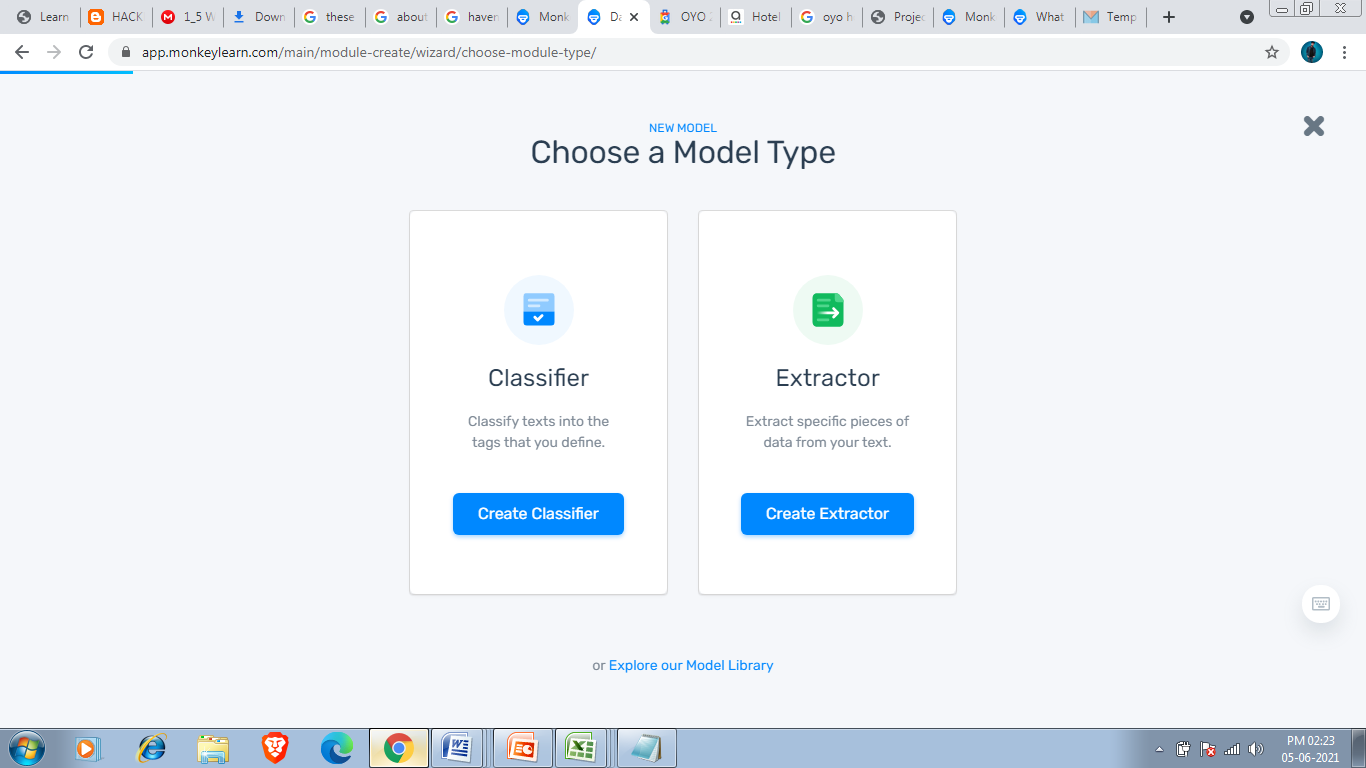
1. **Create a New Classifier**

First of all we have to go to monkey learn website <https://monkeylearn.com/> and you have to make your account after making your account u saw interface as shown in below image.

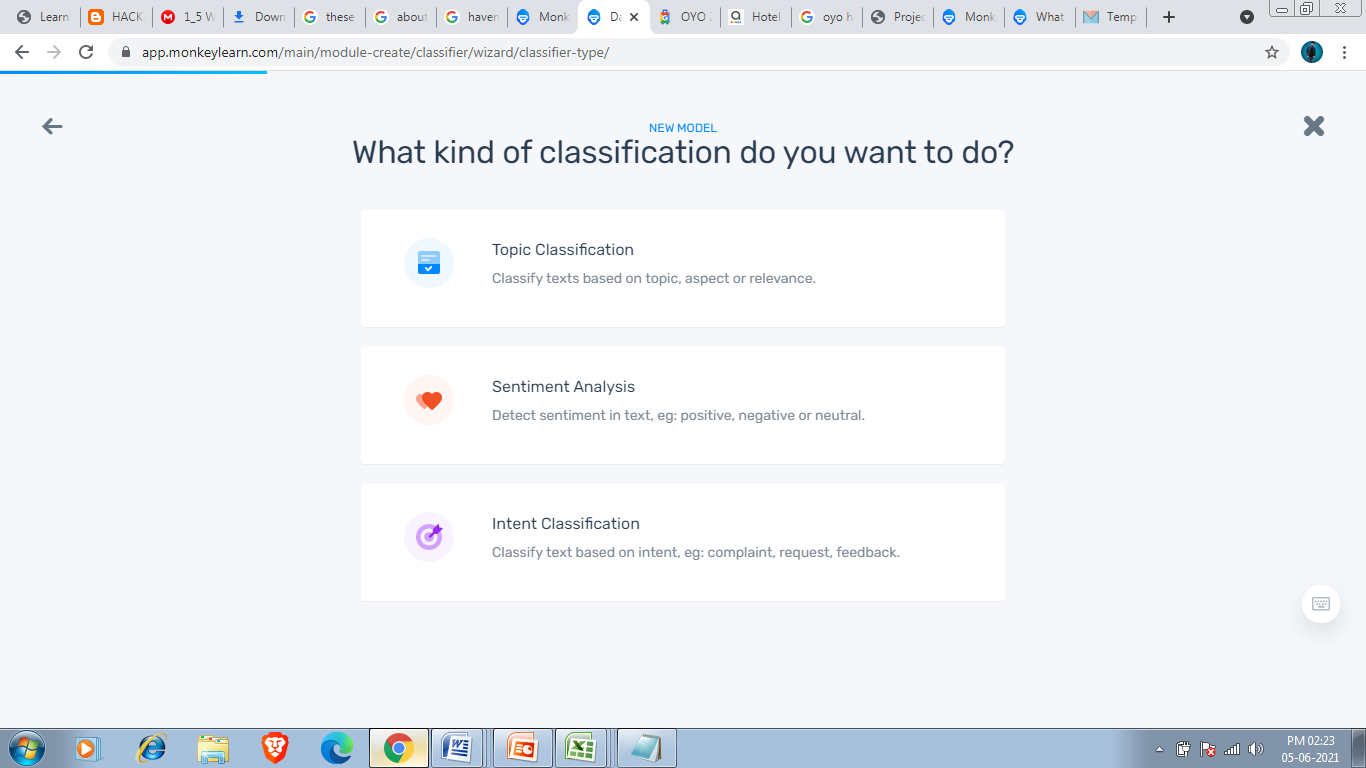


Go to the MonkeyLearn [Dashboard](https://app.monkeylearn.com/main/dashboard/) and click on [Create Model](https://app.monkeylearn.com/main/module-create/wizard/choose-module-type/), then choose Classifier:

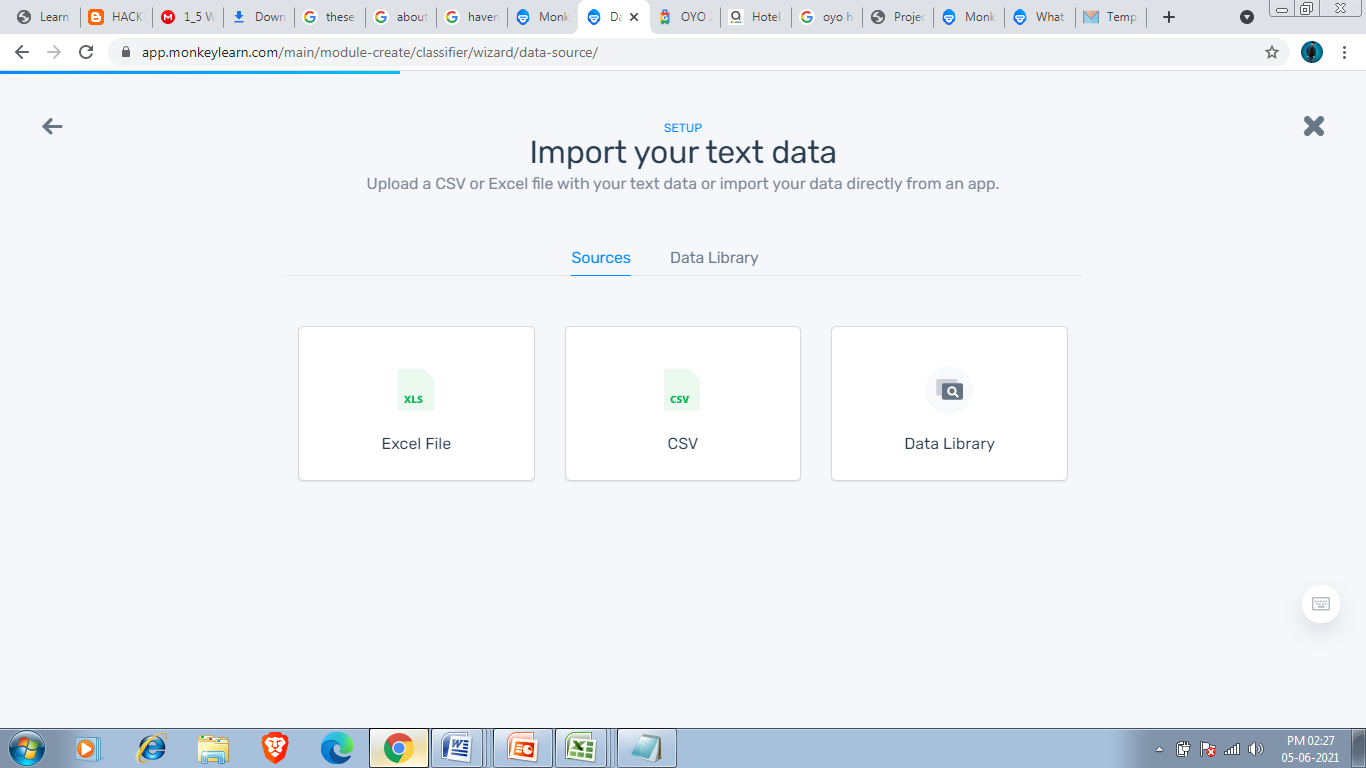
Or click on the blue button that u see in right side.

After Clicking On create model You see menu as shown below image. I have to choose classifier .Because sentimental Analysis is Based on classification. 

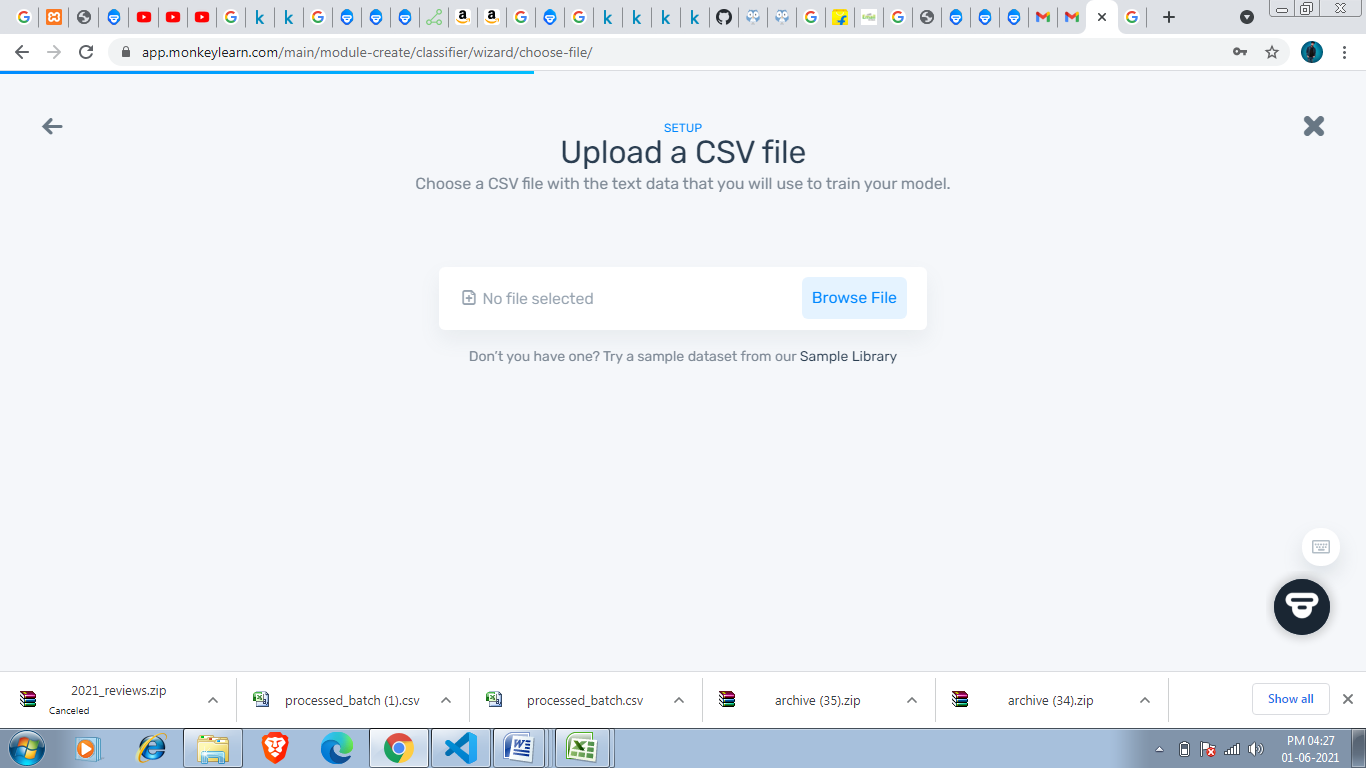
After clicking on classifier you saw 3 options Topic Classification, Sentimental Analysis and Intent Classification .I have to choose Sentimental Analysis Bacause that is my task.



After Clicking on Sentimental Analysis You saw Import your text data option and Sources should we excel or csv file .

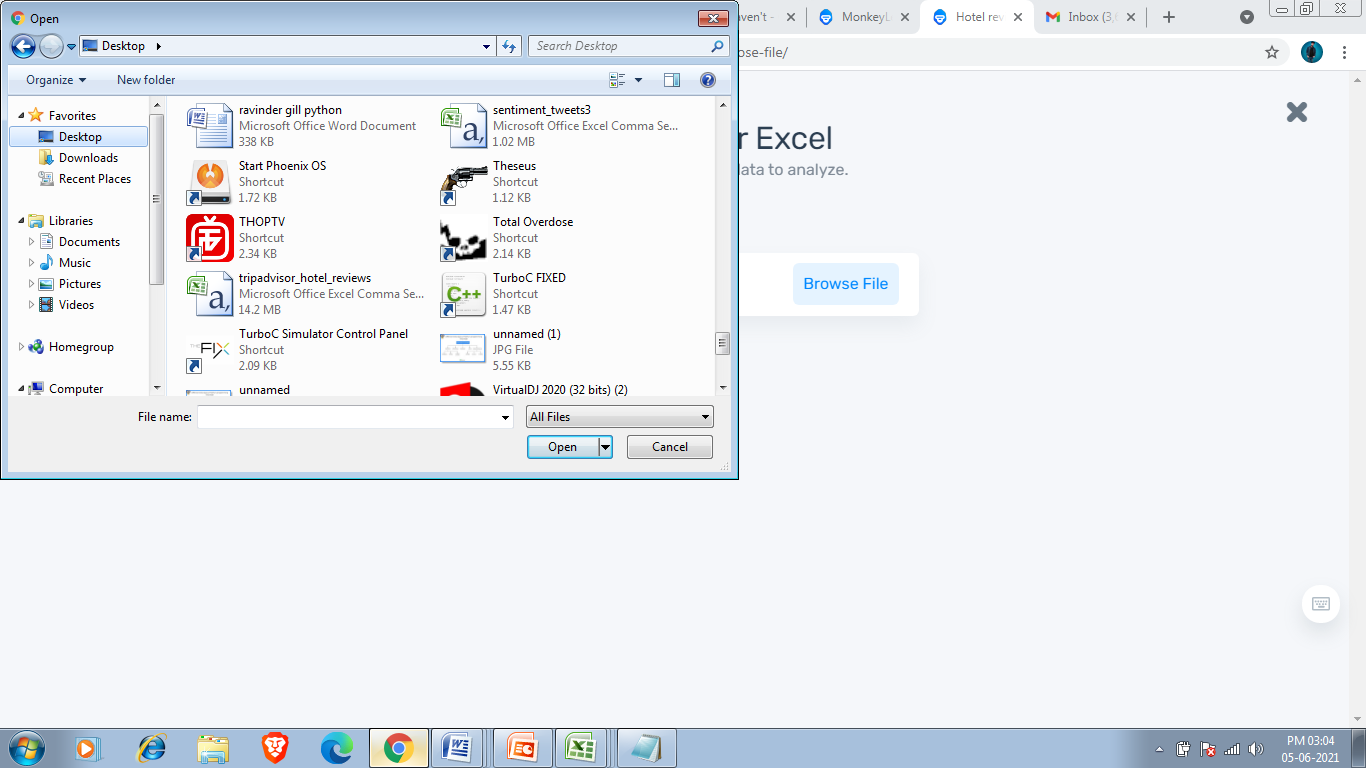


Since my dataset is tripadvisor\_hotel\_reviews.csv which is in csv format so i have to click on csv.

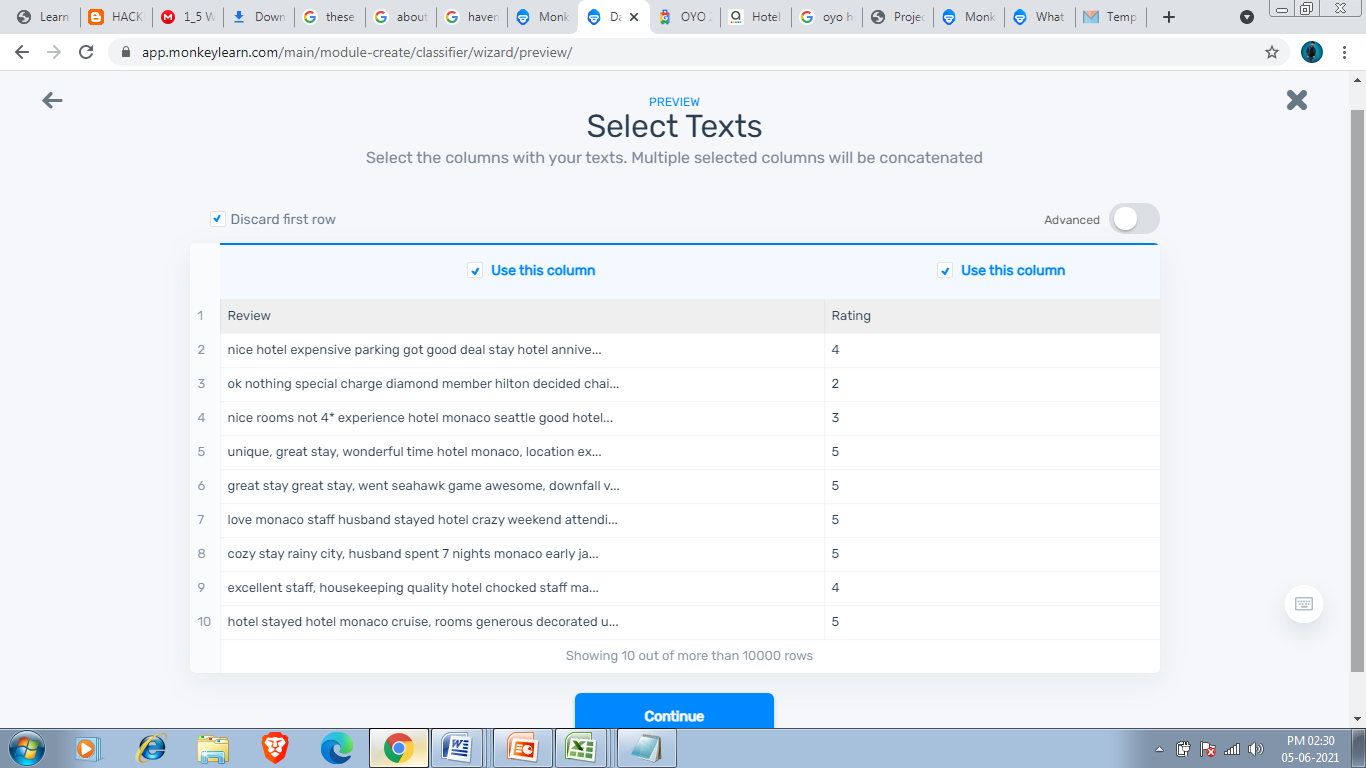


After Clicking on CSV this menu appears where i have to choose my csv by clicking on Browse file my File Explorer opens and i have to choose my data set

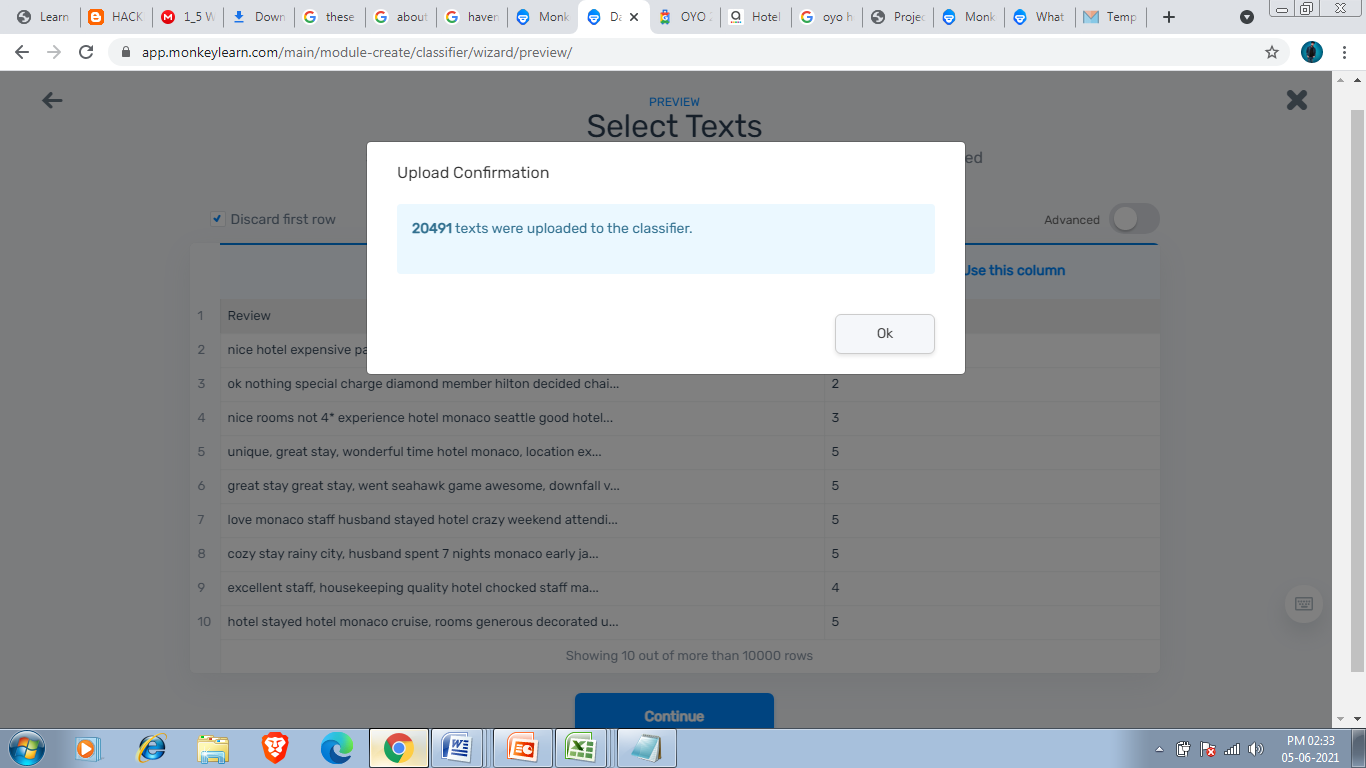
Required data set



After that you have to select columns that you want since i required all columns so selected all after that you have to upload the data by clicking on continue button.

****

Upload confirmation is shown below.

****

After that i see the below menu where i have to define tag.

One of the most critical aspects to building a classifier is defining the tags that you will want to classify for.

The risks of **trying to do too much too fast will have a negative impact** on the performance of a model. Though machine learning models can eventually be trained to be complex things, all of them had to start doing something simple first.

## Sets of Tags and Examples

Some tags are pretty self explanatory. For example sentiment analysis generally includes:

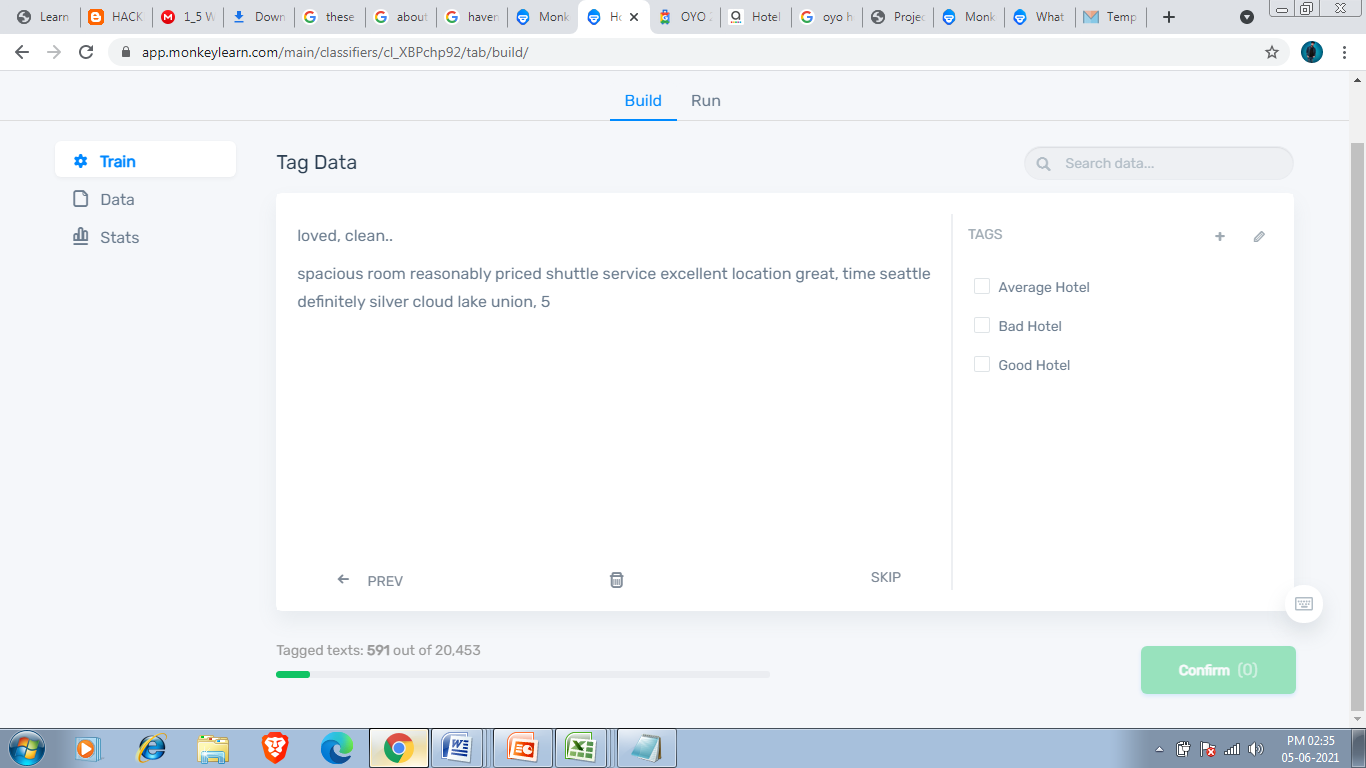
* Positive
* Negative
* Neutral

So i am defining 3 Tags:-

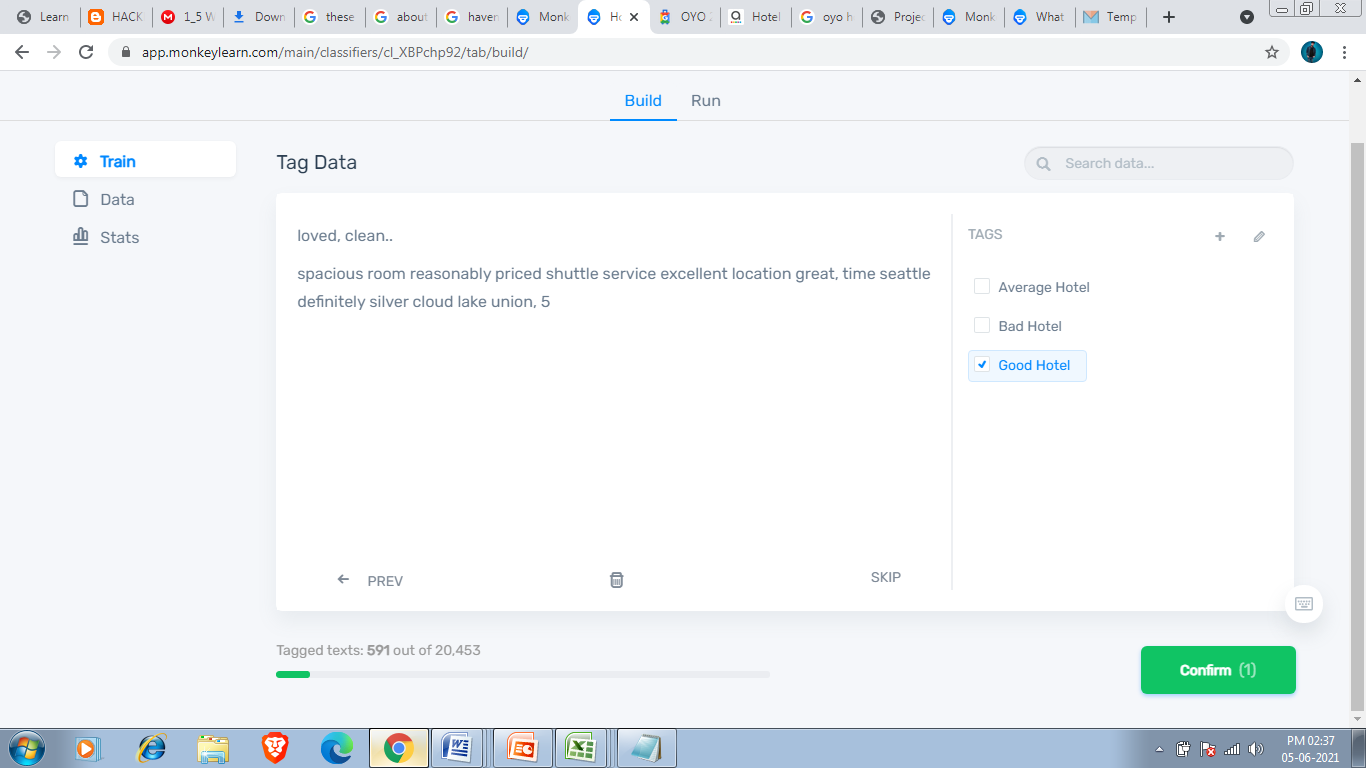
Good Hotel

Bad Hotel

Average Hotel



So i have defined 3 tags as shown in above figure after this i saw menu where we have to tag data .Here We see reviews of different users and we have to tag according to the rating.

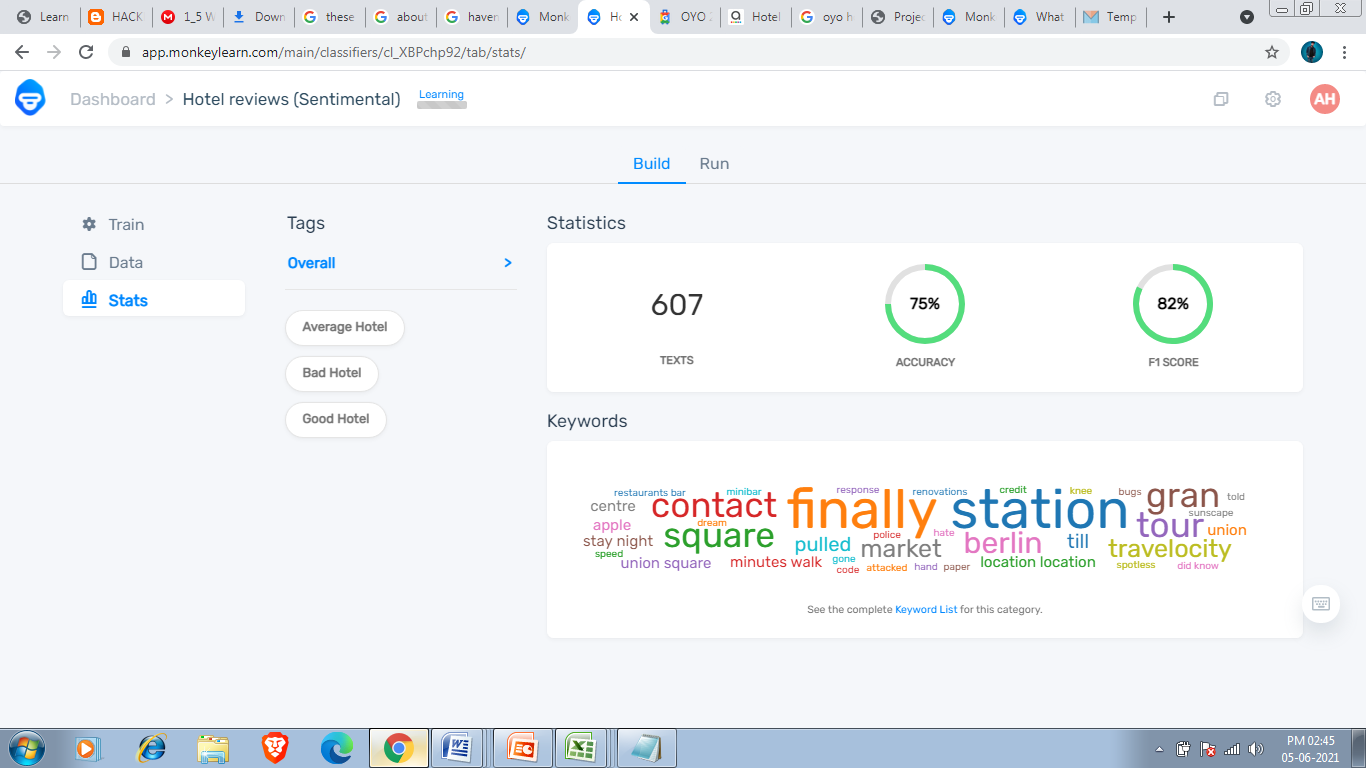


We saw from above image that customer is happy from hotel service so according to him hotel is good so i am tagging this to Good Hotel Tag.

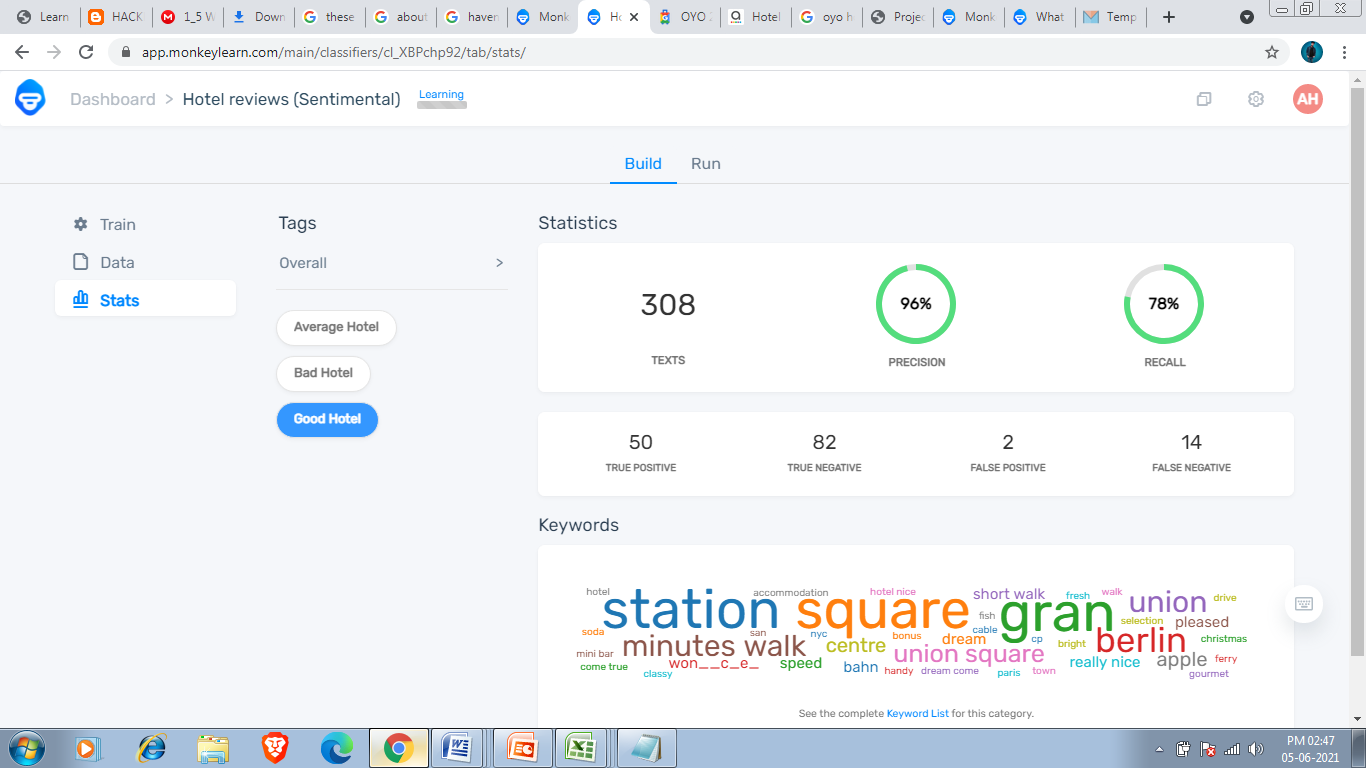
After clicking on confirm button this review is get tagged to Positive tag.we have to Tag atleast 20 But the more you tag the more accurate is your model .

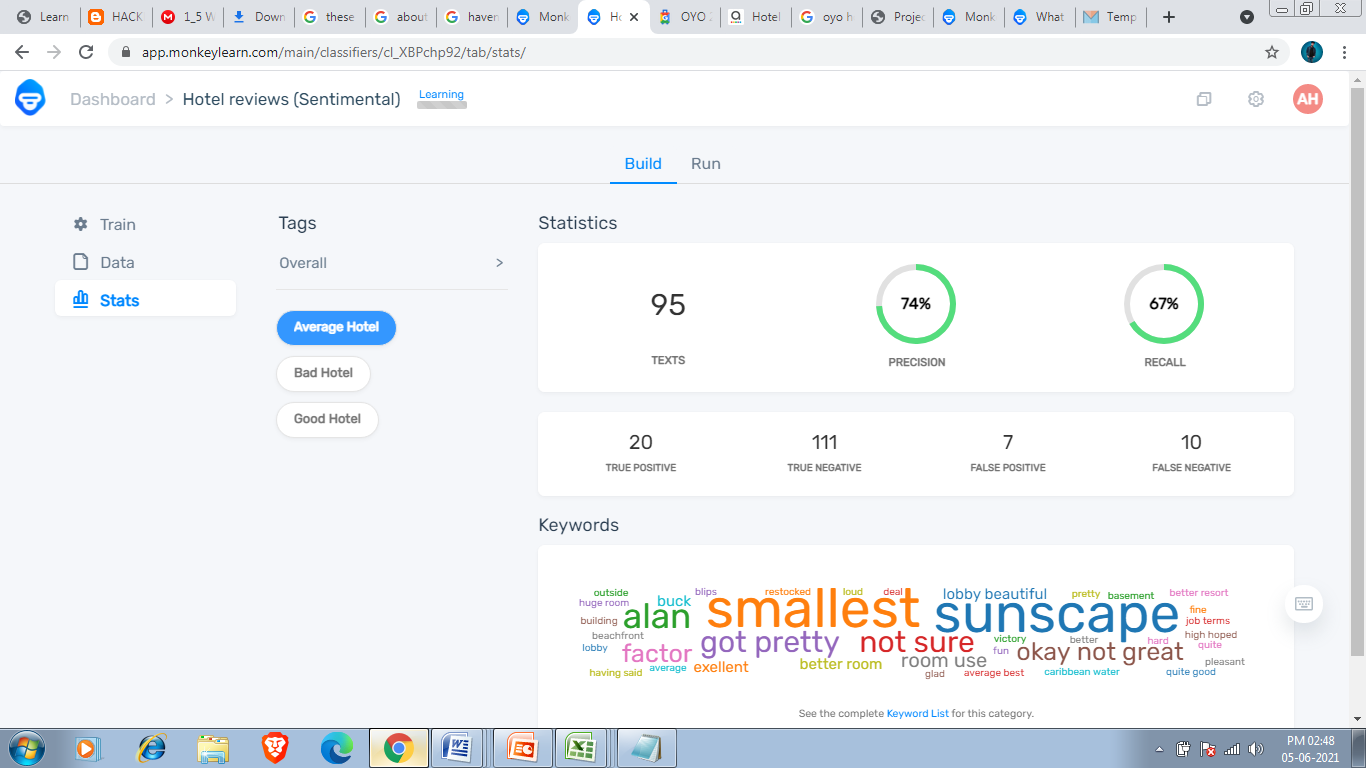
So i tagged 591 customer reviews.We tag our data to train its so that it give us accurate result.

My model’s overall accuracy is 75%.

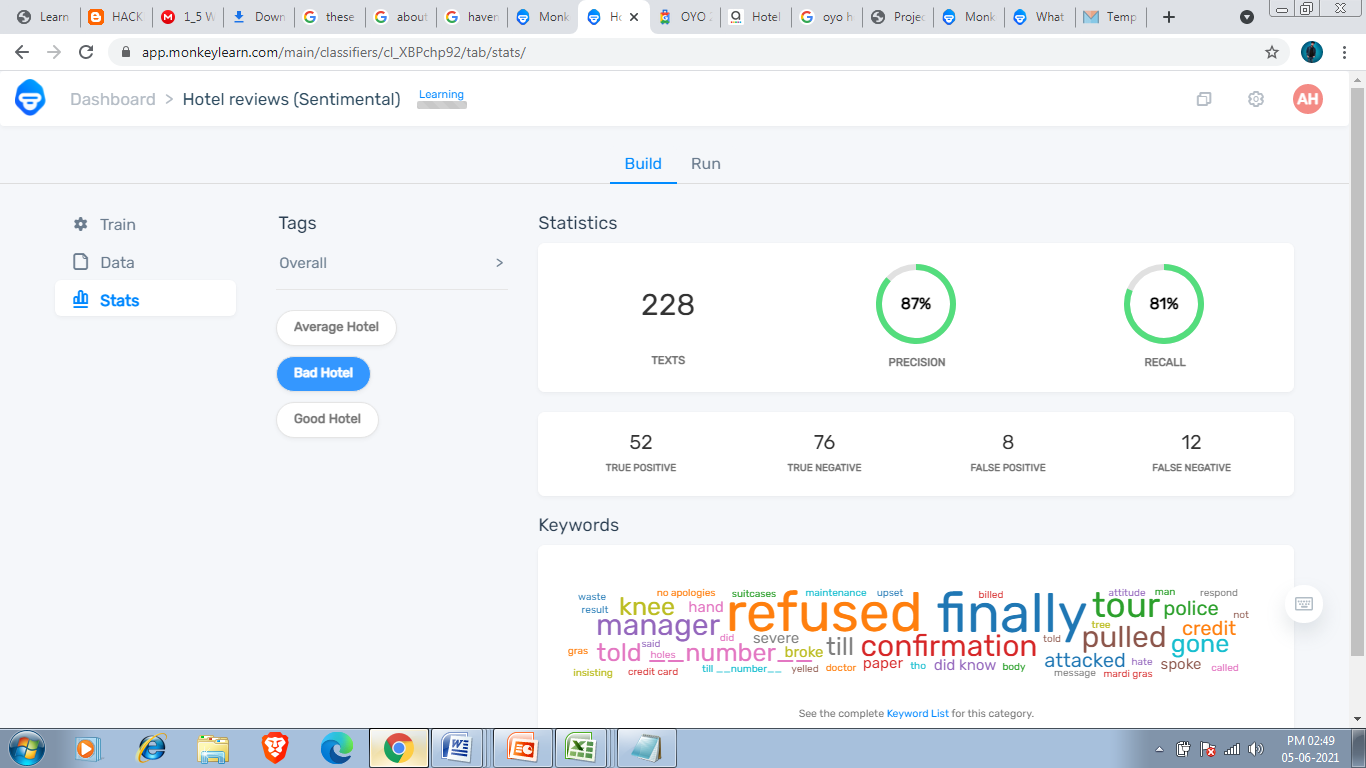


My models’s Good Hotel tag is 96% precision and 78% recall.



My models’s Average Hotel tag is 96% precision and 78% recall. 

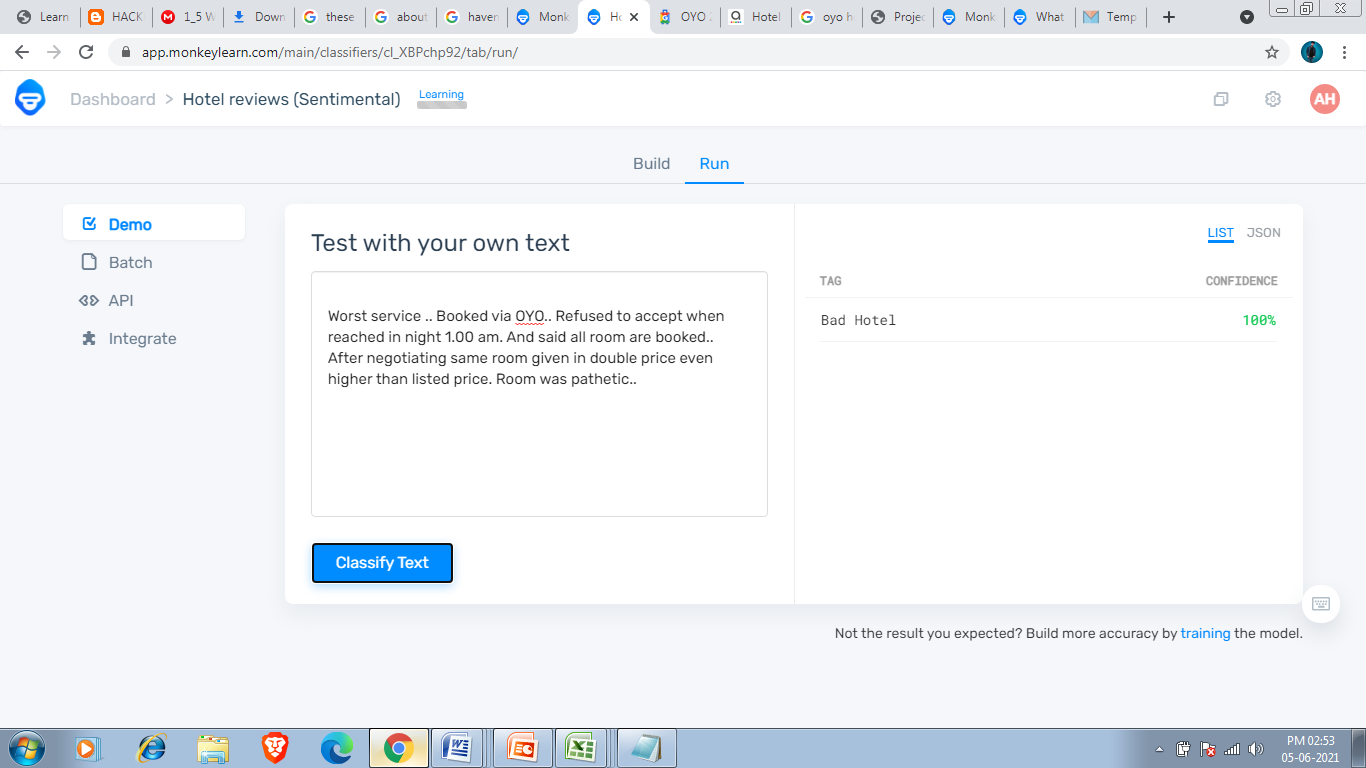
My models’s Bad Hotel tag is 87% precision and 81% recall.



I can increase my model accuracy by proper tagging.

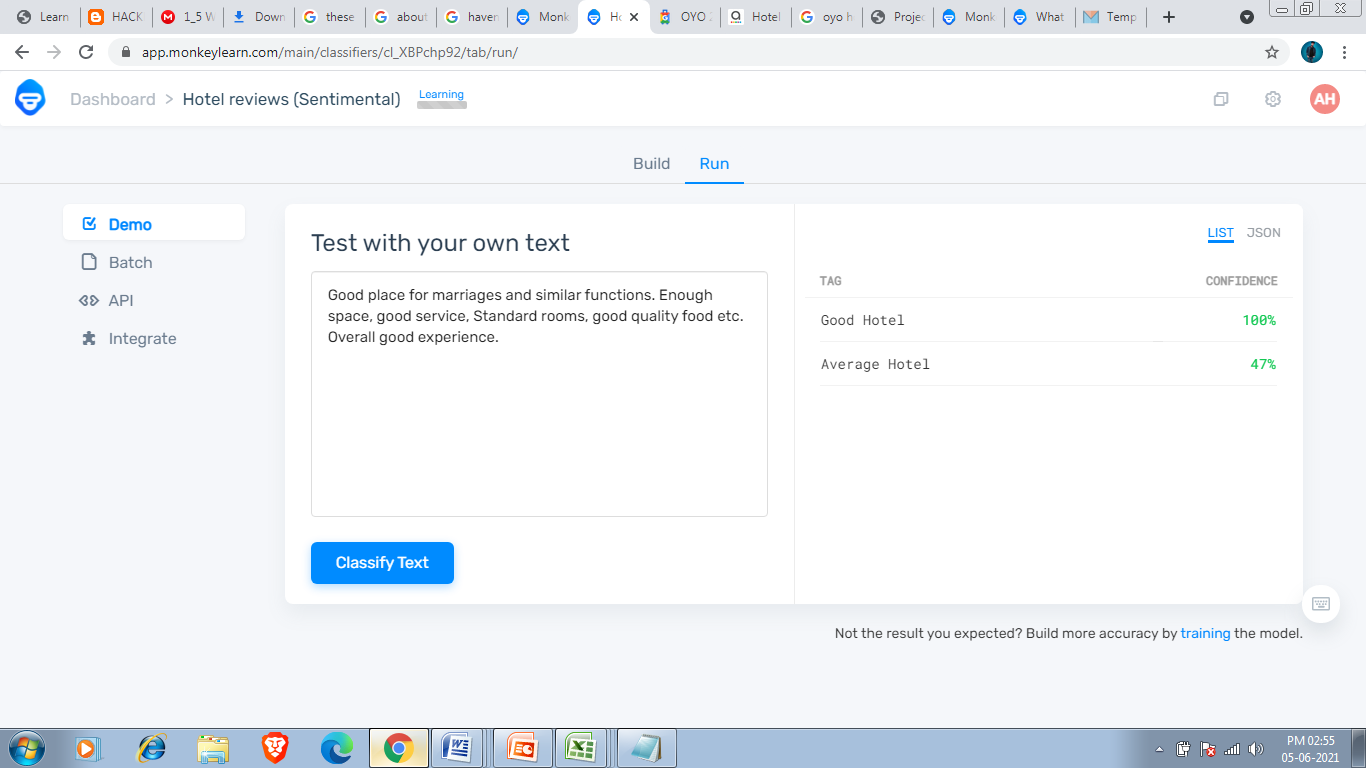
We can test our model by batch file and demo.

In demo we write text such as review of hotel and it will tell this review fall in which category i am taking reviews of my near hotels for demo.



So according to my model person is saying hotel is bad and that is correct.

Testing another text.



So according to my model this review has 100% confidence means person have good opinion towards hotel and that is correct.

I can also process my file and after processing the resultant file tell about the review either hotel good,bad or average and how much confidence it have.