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

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.

Much of the documentation and terminology in our platform will refer to classification models as classifiers, and to extraction models as extractors.

So in this assignment i am performing sentimental analysis of AllProductsReviews.csv data set.

This data set contains Amazon reviews of 10 earphones ( boAt Rockerz 255, Flybot Wave, Flybot Boom ,PTron Intunes, Flybot Beat, Samsung EO-BG950CBEIN, JBL T205BT, Sennheiser CX 6.0BT, Skullcandy S2PGHW-174 and JBL T110BT).These data sets are created by Scrapping so i am still learning it so i have to choose data that is scrapped by other person. I am creating a model that will help me to know about people opinion about these products.

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/), 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)

So How To Do this Assignment .I explained it below

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 in this sentimental analysis i am using AllProductsReviews.csv which i downloaded from kaggle.

This data set has 14337 Reviews of 10 Products.

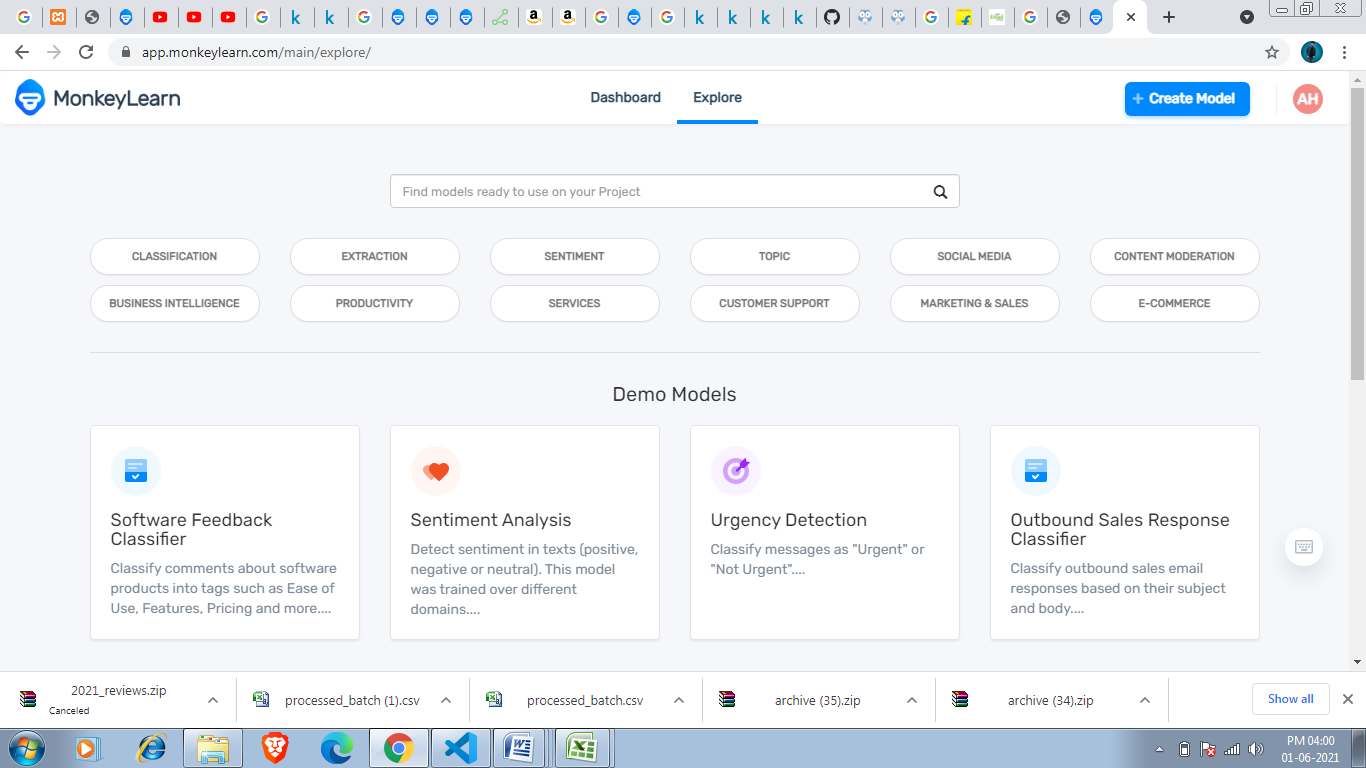
Because i have not much knowledge about scrapping i use my own scrapped data in my further projects.

**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**

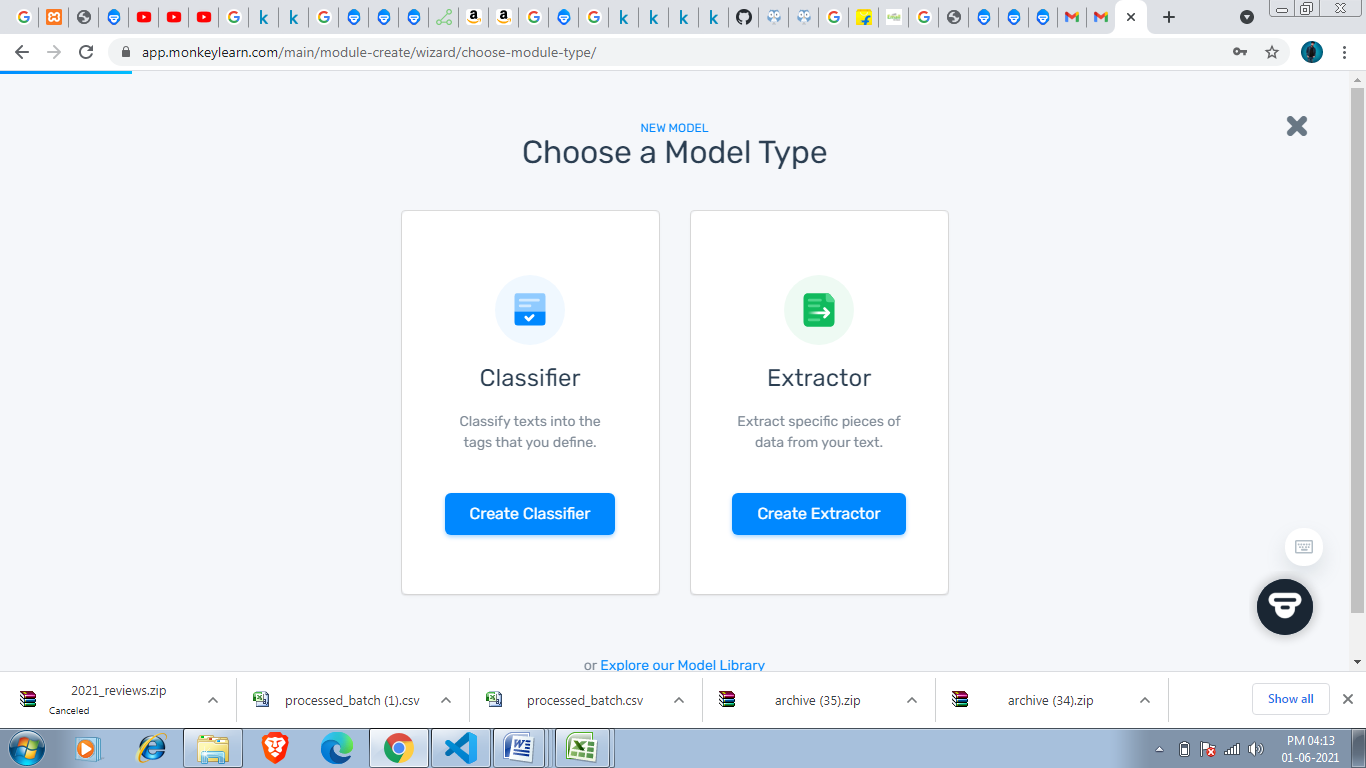
First of all we have to go to monkey learn website <https://monkeylearn.com/> and you have to make your account after making your account u saw interface as shown in below image.



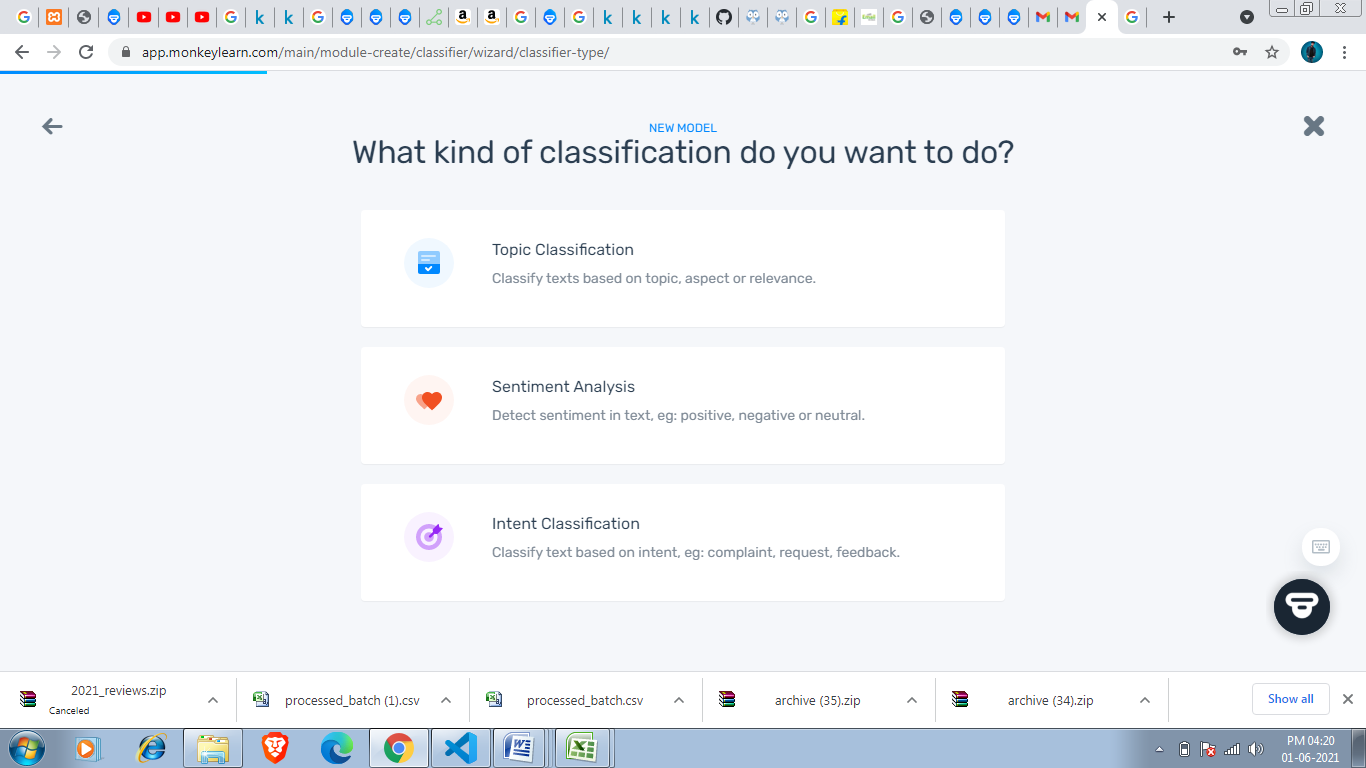
Go to the MonkeyLearn [Dashboard](https://app.monkeylearn.com/main/dashboard/) and click on [Create Model](https://app.monkeylearn.com/main/module-create/wizard/choose-module-type/), then choose Classifier:

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

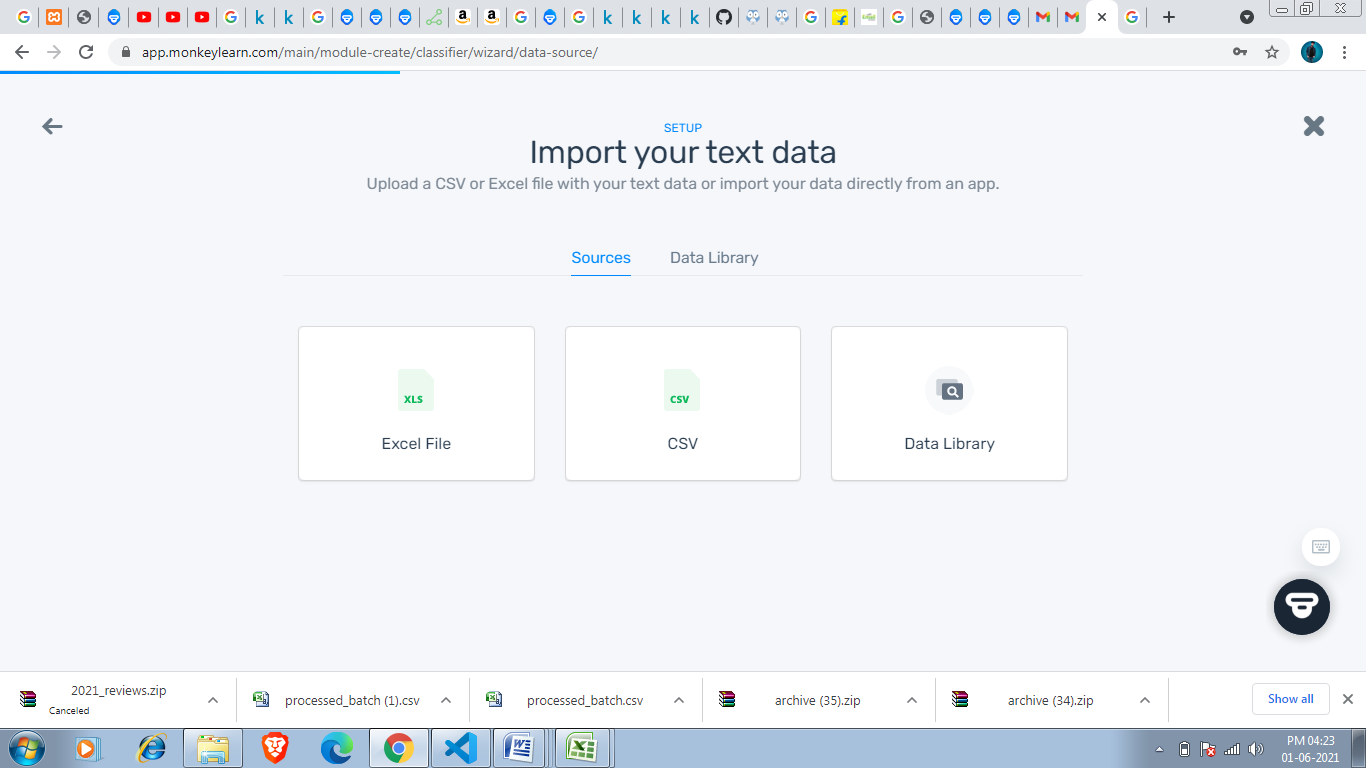
After Clicking On create model You see menu as shown below image. I have to choose classifier .Because sentimental Analysis is Based on classification.

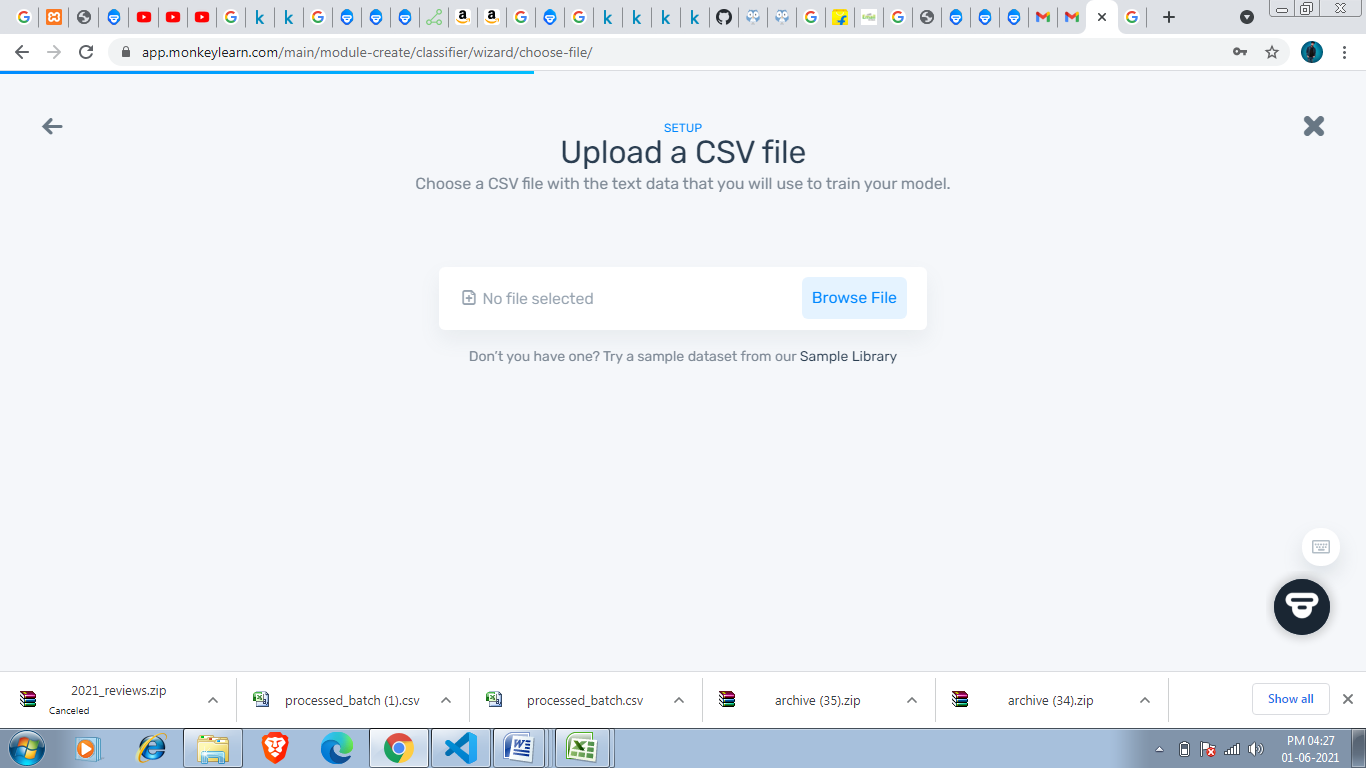


After clicking on classifier you saw 3 options Topic Classification, Sentimental Analysis and Intent Classification .I have to choose Sentimental Analysis Bacause that is my task.



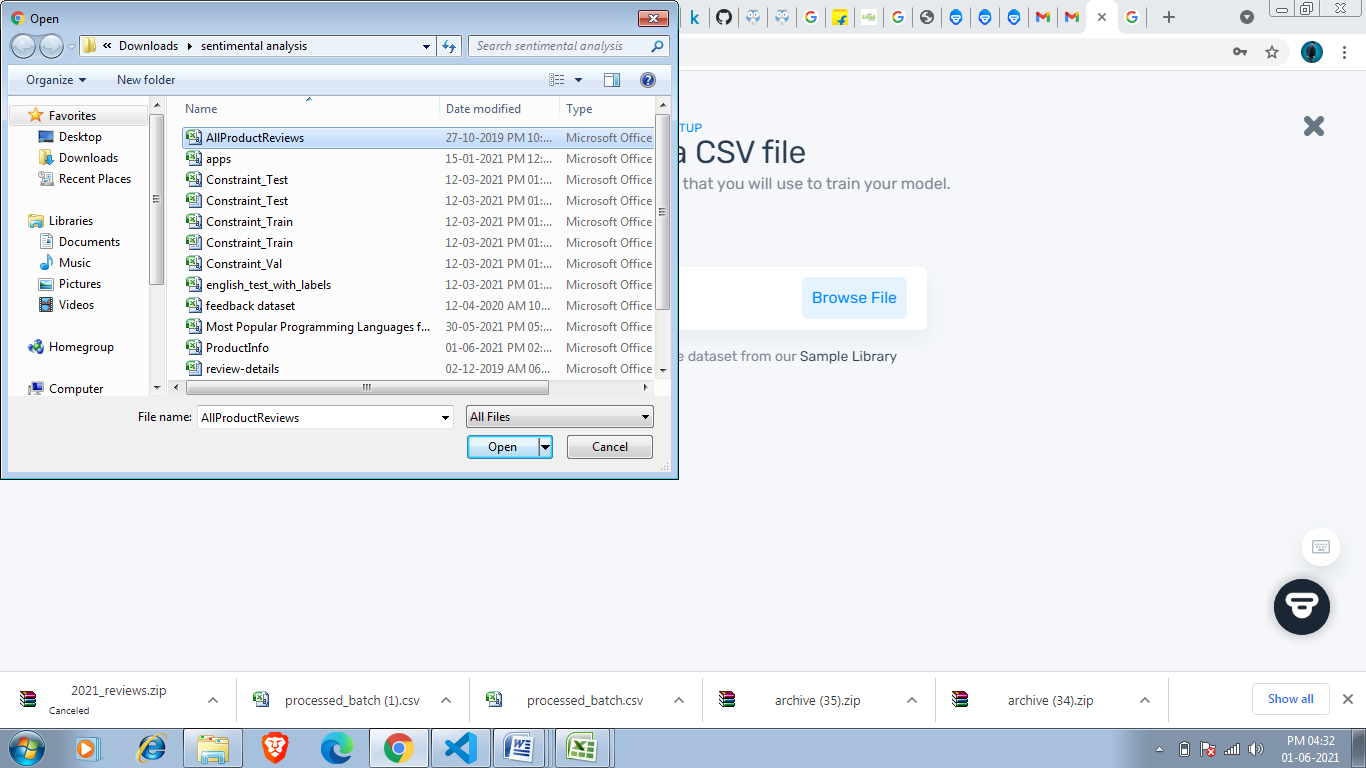
After Clicking on Sentimental Analysis You saw Import your text data option and Sources should we excel or csv file .



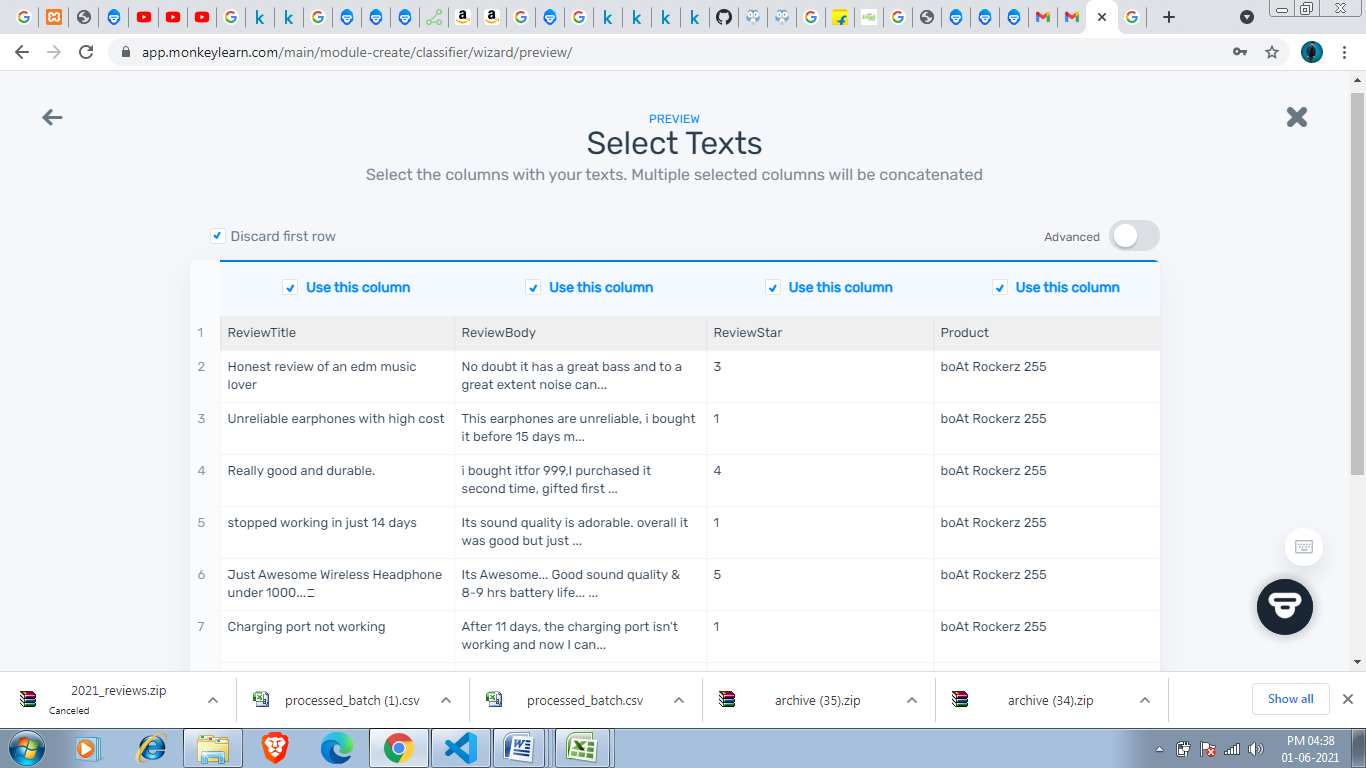
Since my dataset is **AllProductsReviews.csv** which is in csv format so i have to click oncsv. 

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.

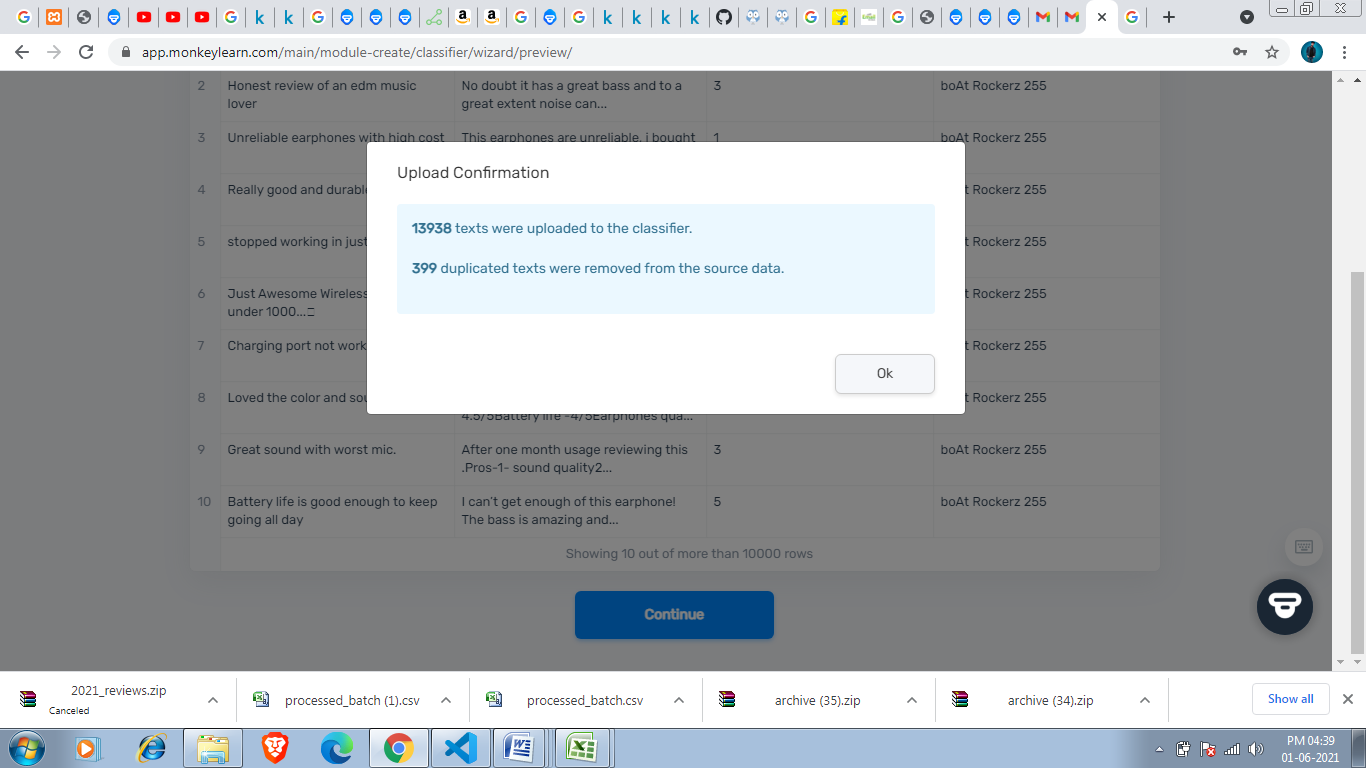
Required data set



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.



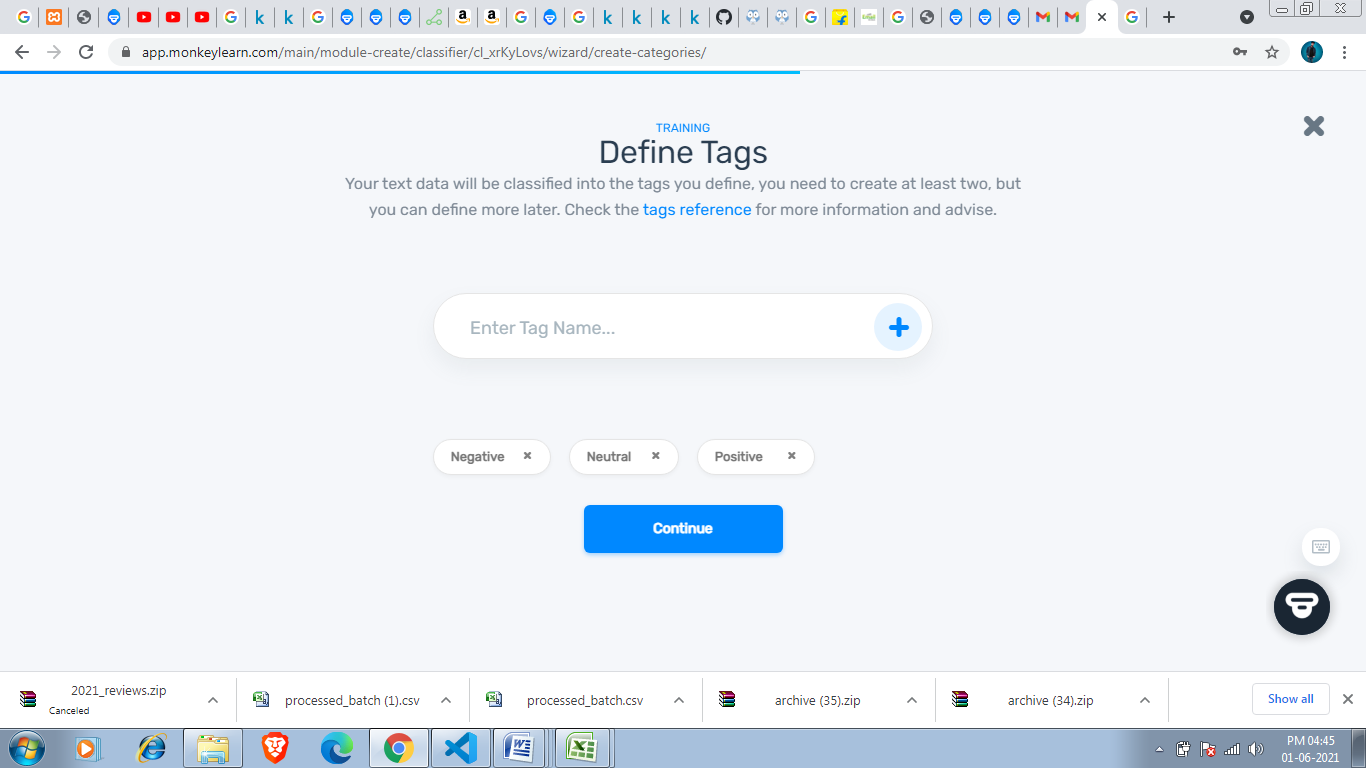
Upload confirmation is shown below.



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:

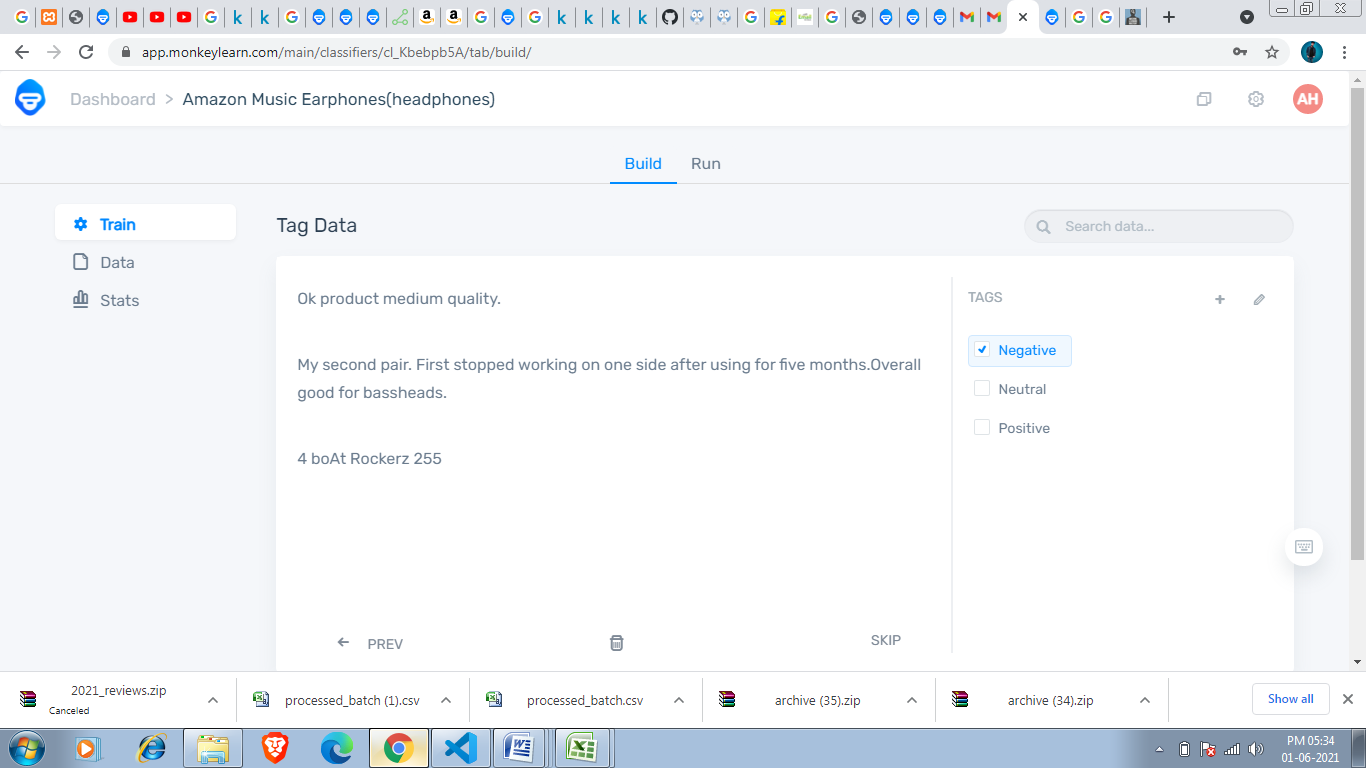
* Positive
* Negative
* Neutral

So i am defining 3 Tags

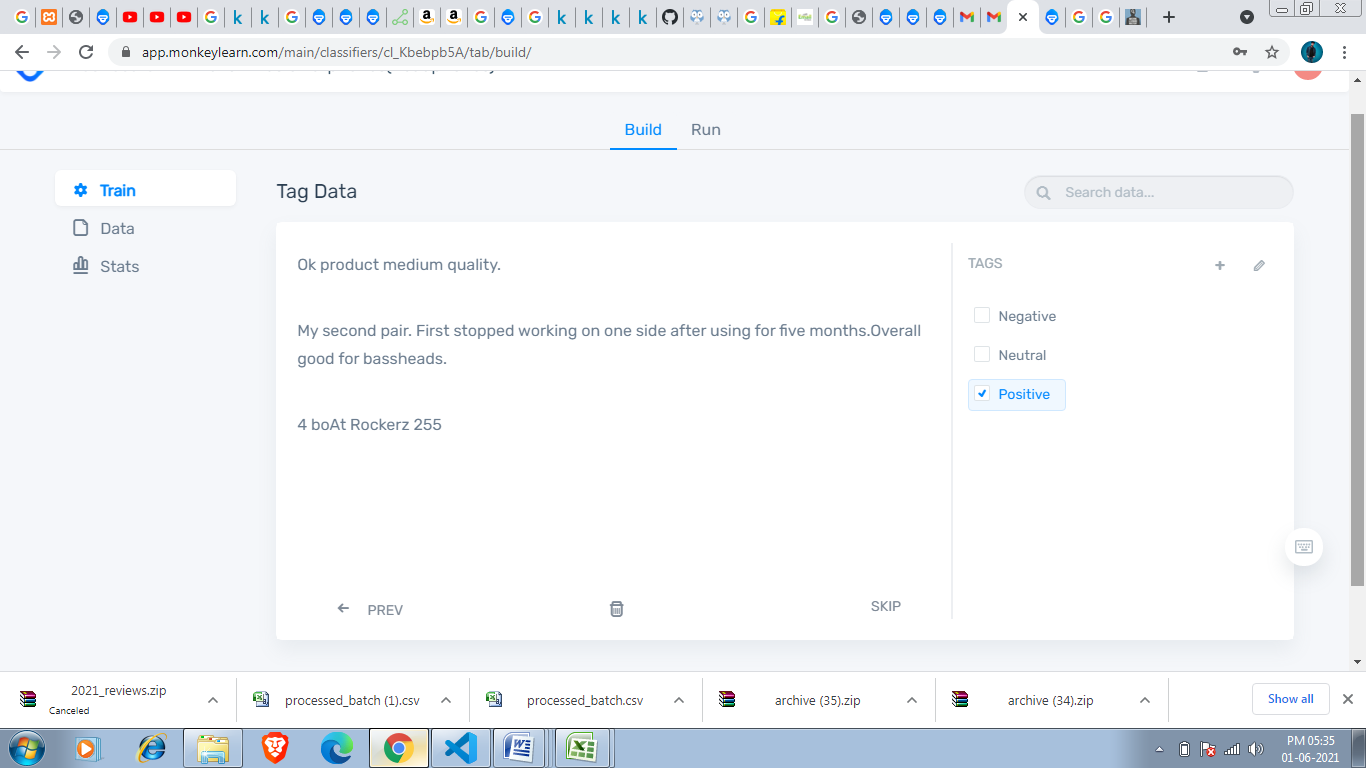
Positive

Negative

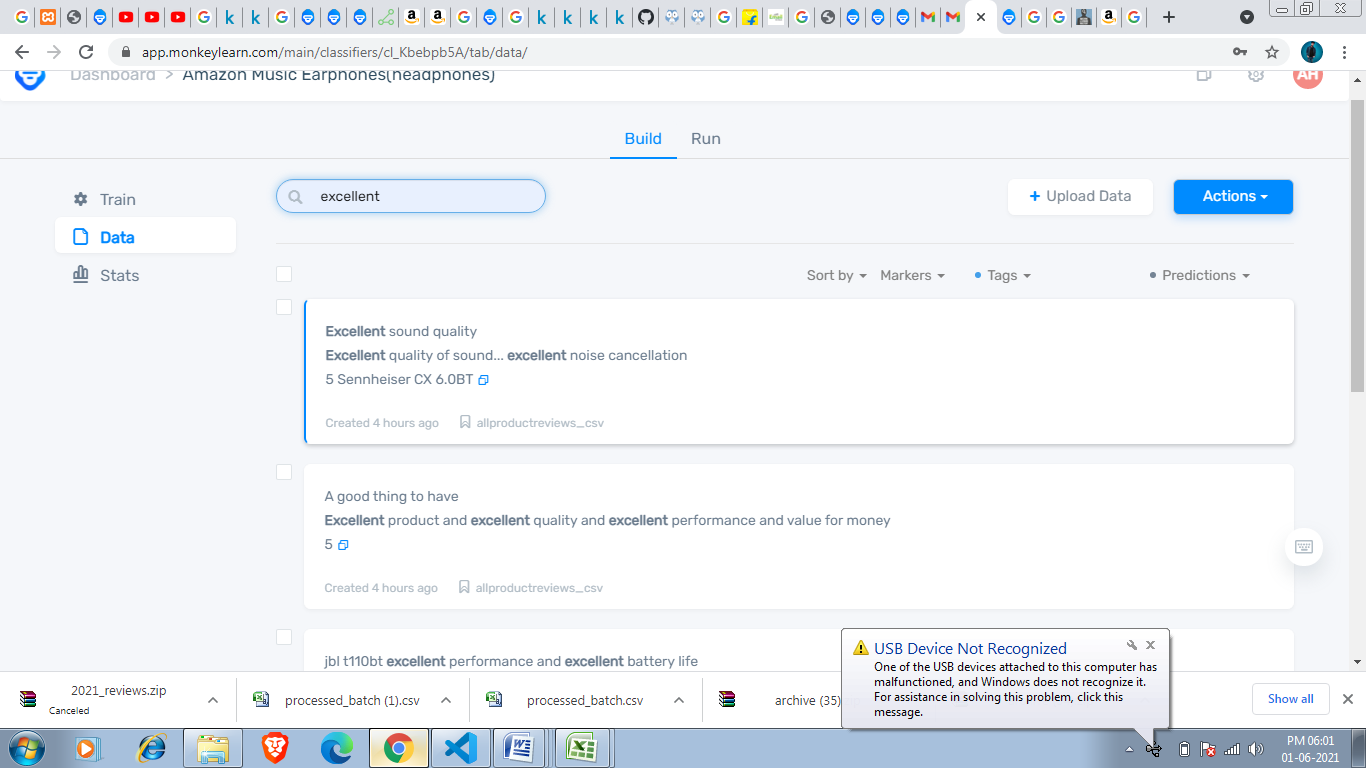
Neutral



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.

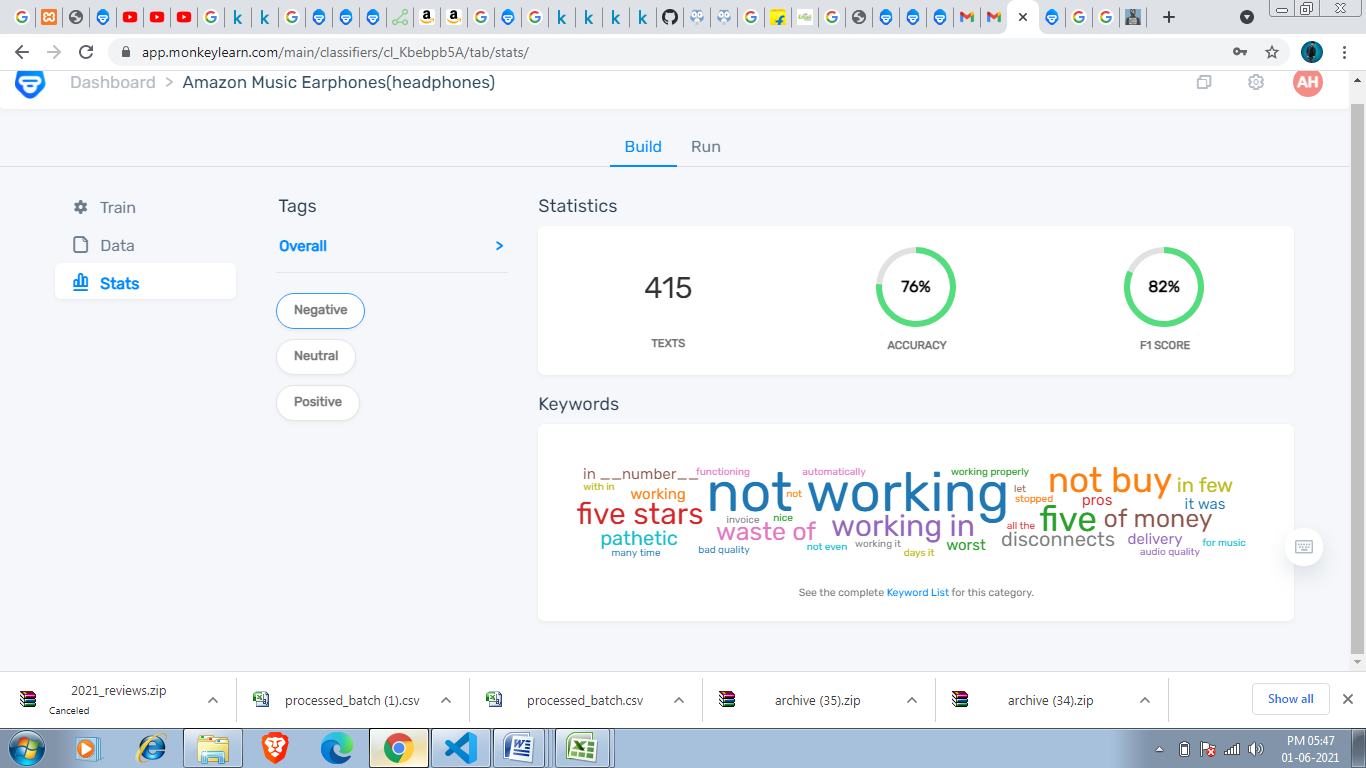


After That i have to tag data according to the review as in above figure rating is 4 so i can tag it to Positive as shown in below figure.

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 am tagging atleast 200-300 of 13938 Reviews and i can’t show how i tagged these 200-300 entries because its huge number. 

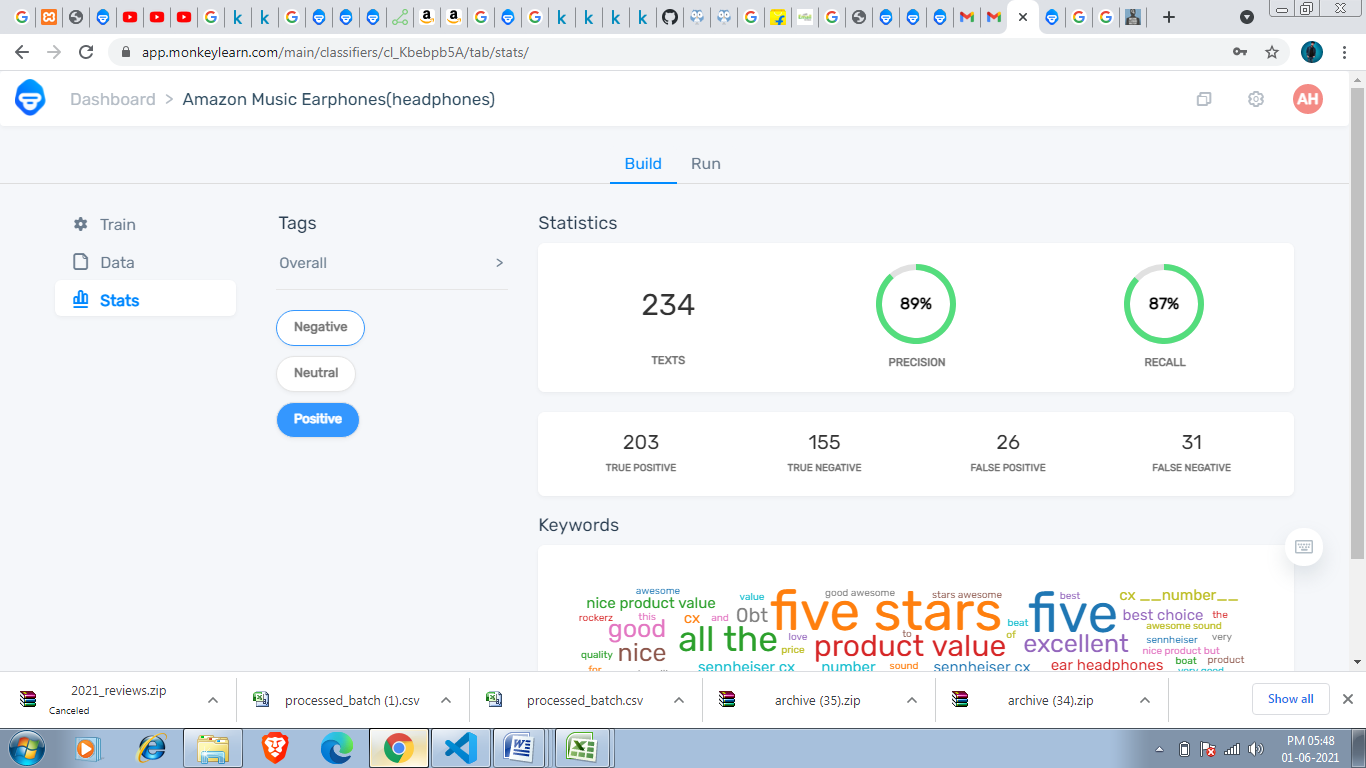
We can search for particular data in Data option we get to know about prediction.

I have tagged 415 entries and my overall model is 76% accurate i can make it more accurate by tagging rightly.It also shows main keywords.



And keywords show most common words

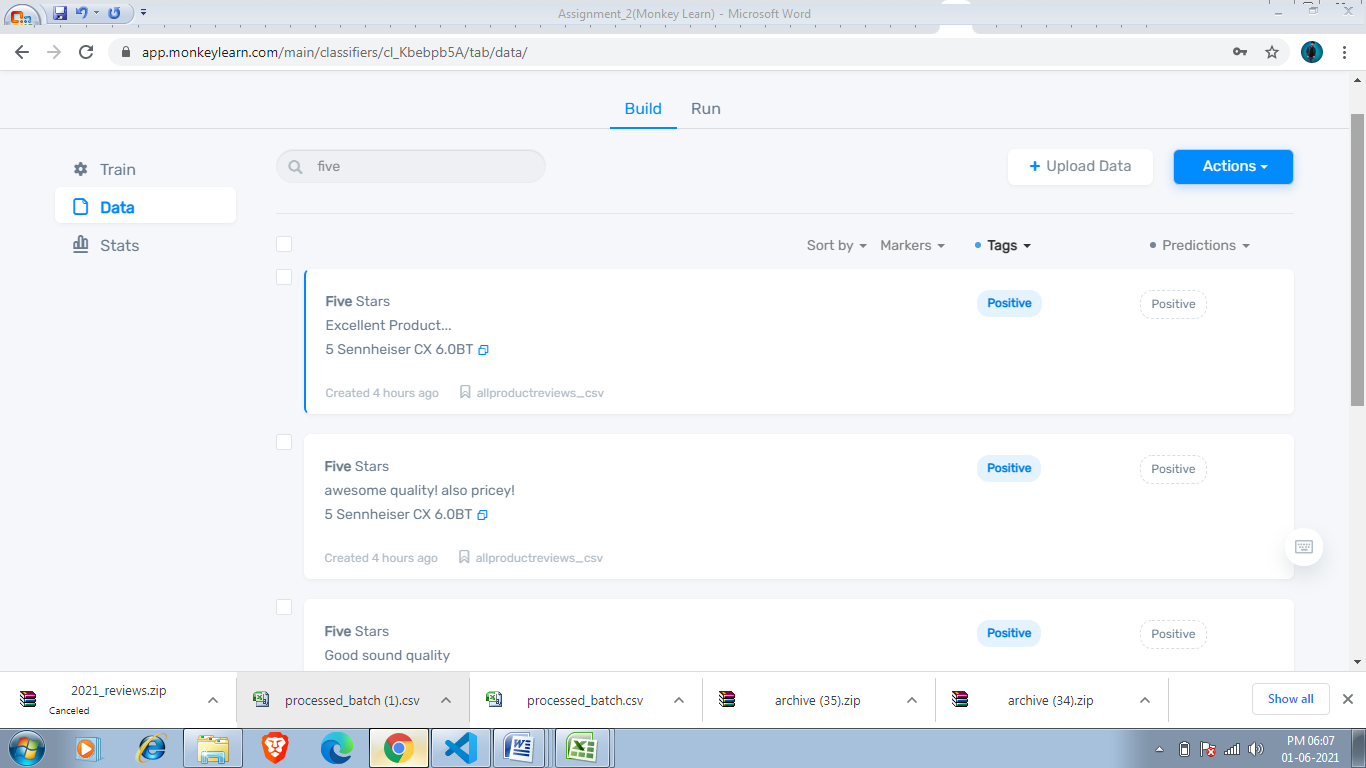
My model’s positive tag is 89% precision.



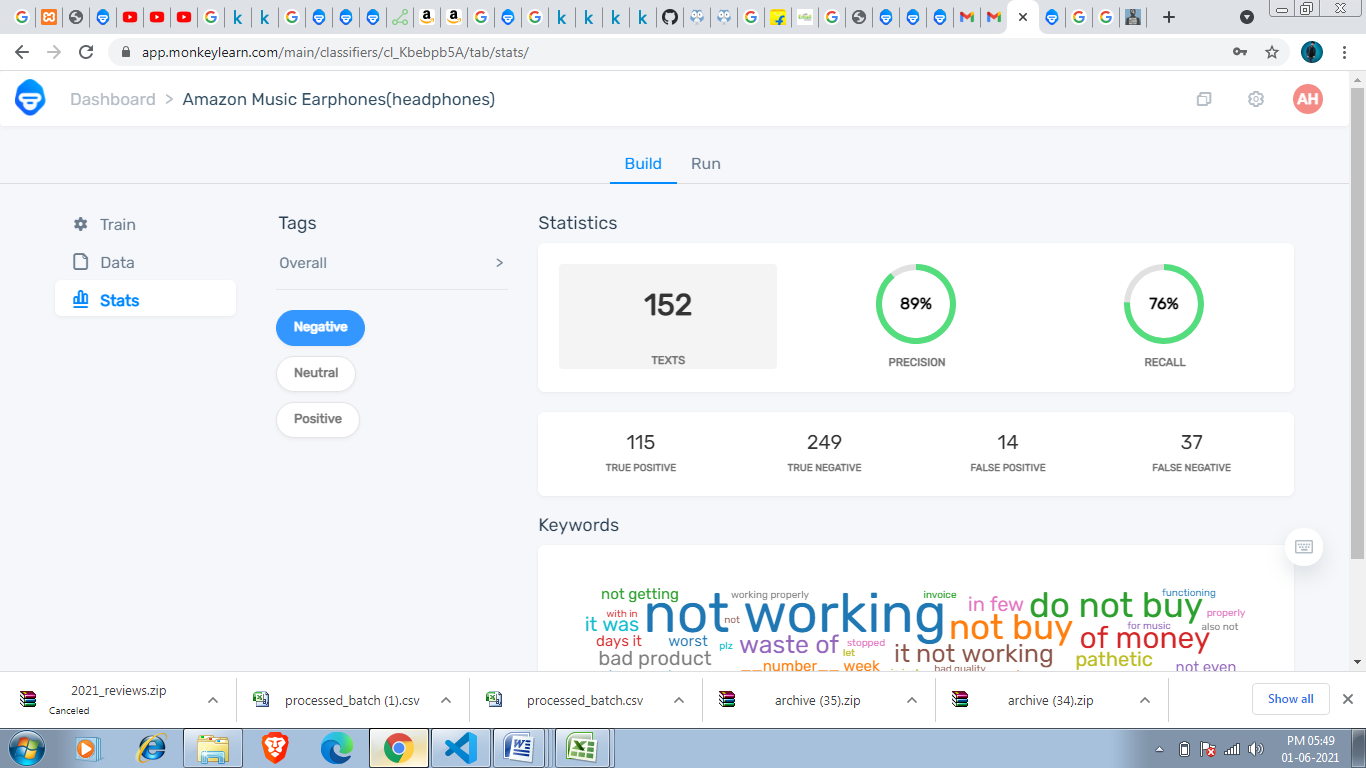
Five star ,five ,excellent are most common word in Positive Tag.

So when i click on these keywords then it show me all prediction related to that keyword.

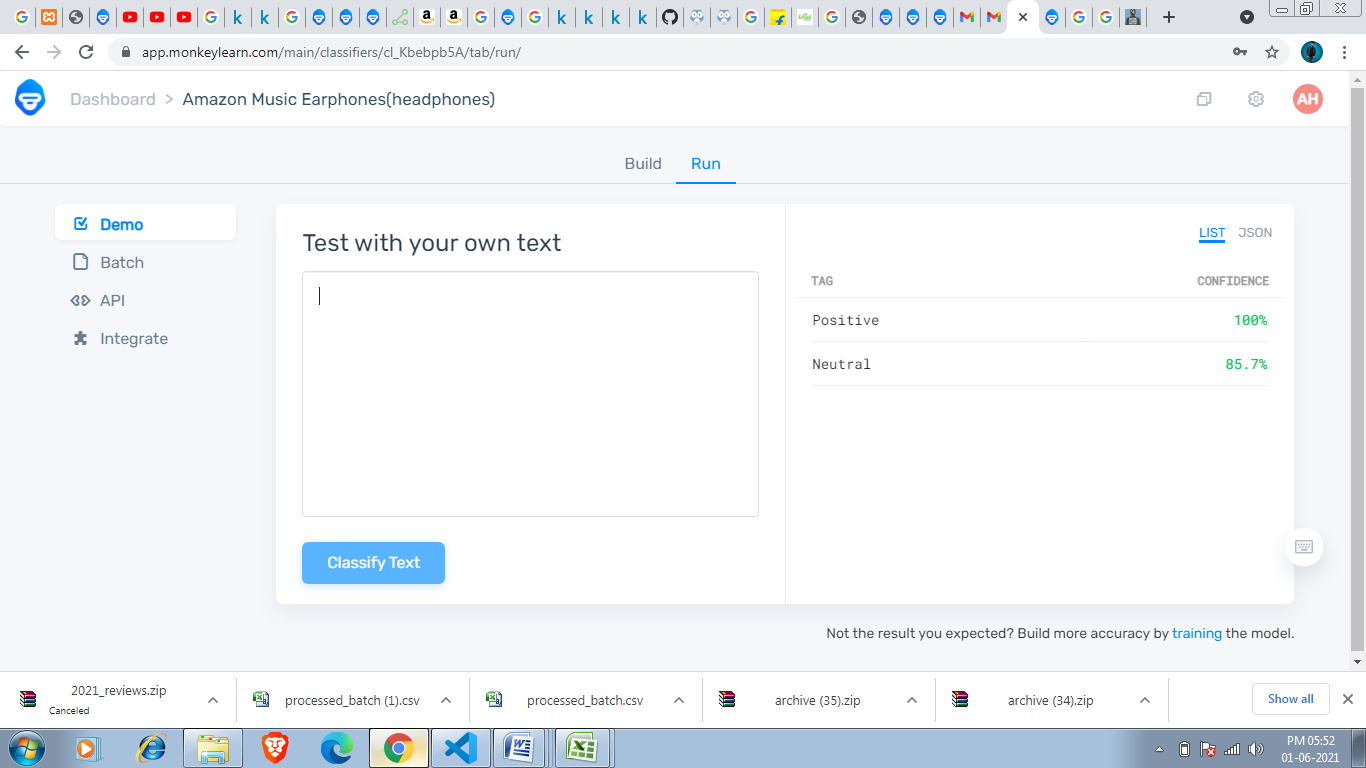
So when i click on five star then i get result as shown below.



My Model’s Negative tag is 89% accurate(précised).

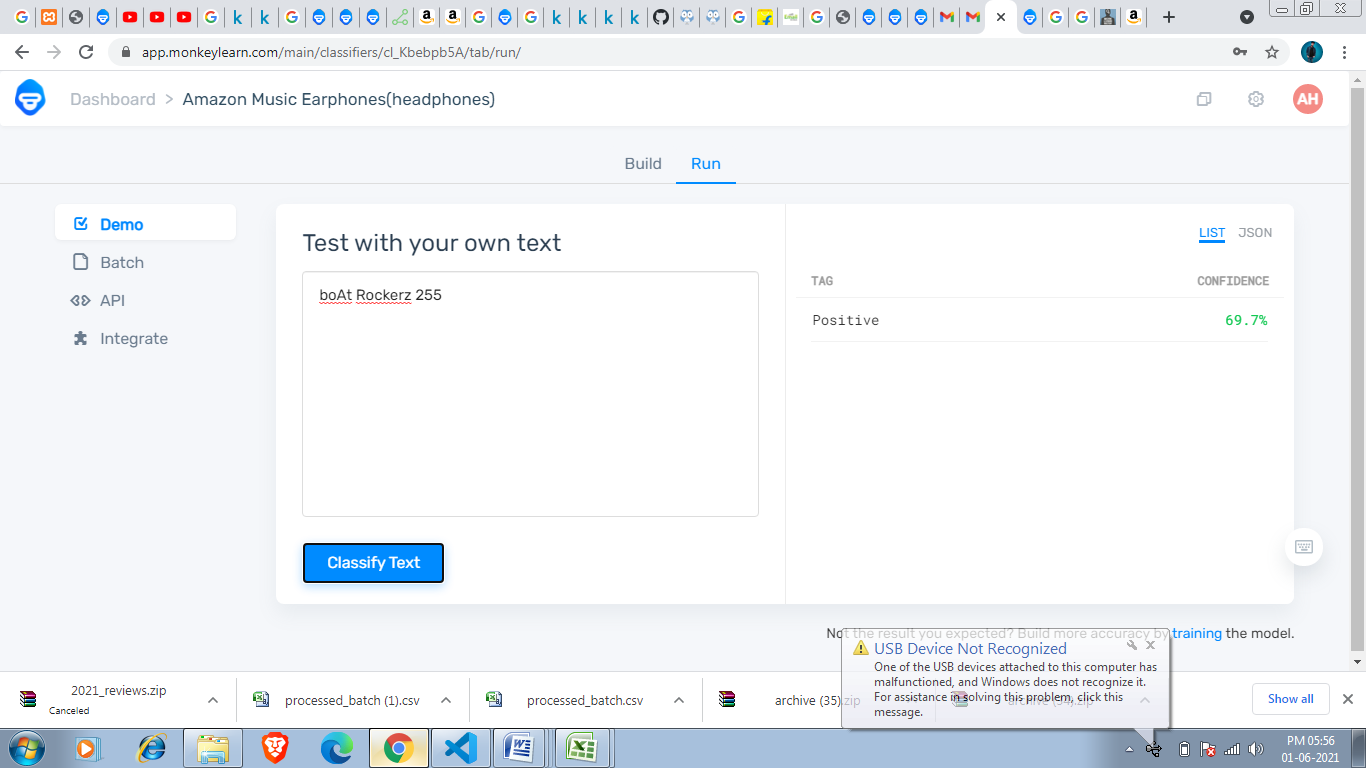


Now i can run my model as shown below



We just have to enter task whose confidence i have to check.

So i am writing boAt Rockerz 255 because i want to check is confidence.

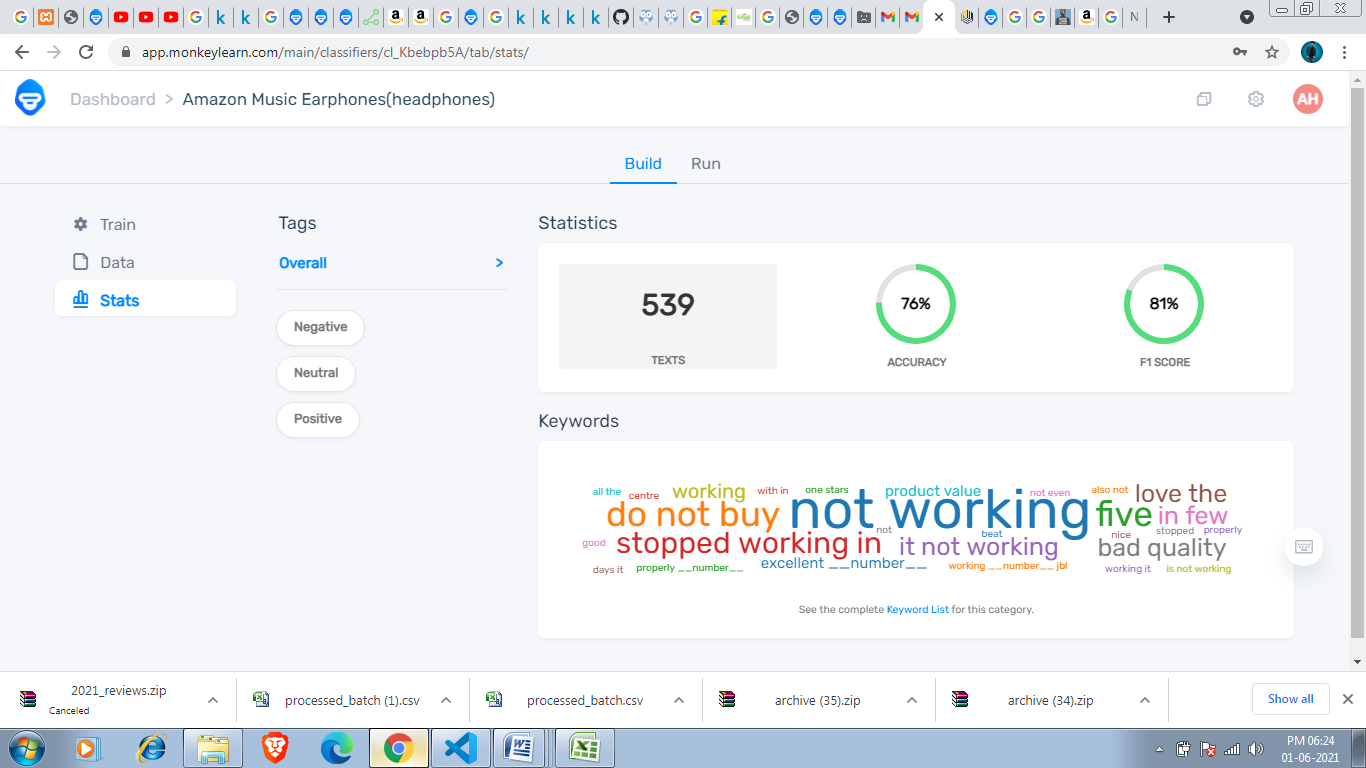


My model shows it have 69.7 % good reviews confidence which means it is a good product.

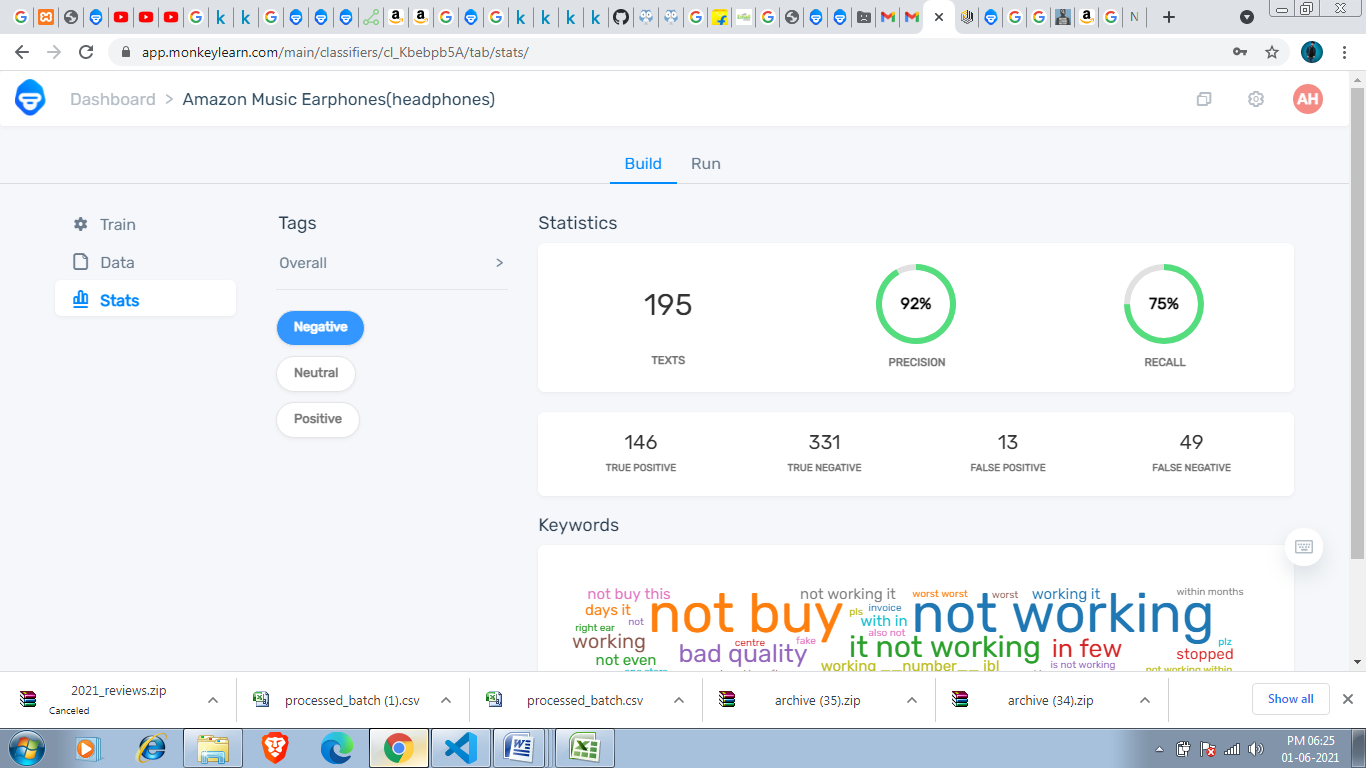
We can also integrate zapier ,Rapid Miner and Google sheets with Monkeylearn .

After tagging 539 reviews my model become 76% accurate ,Negative tag has 92% precision,Positive tag has 87% precision and Neutral has 69% precision.This model help me to know about customer. What is negatives and positives in our product according to customer.

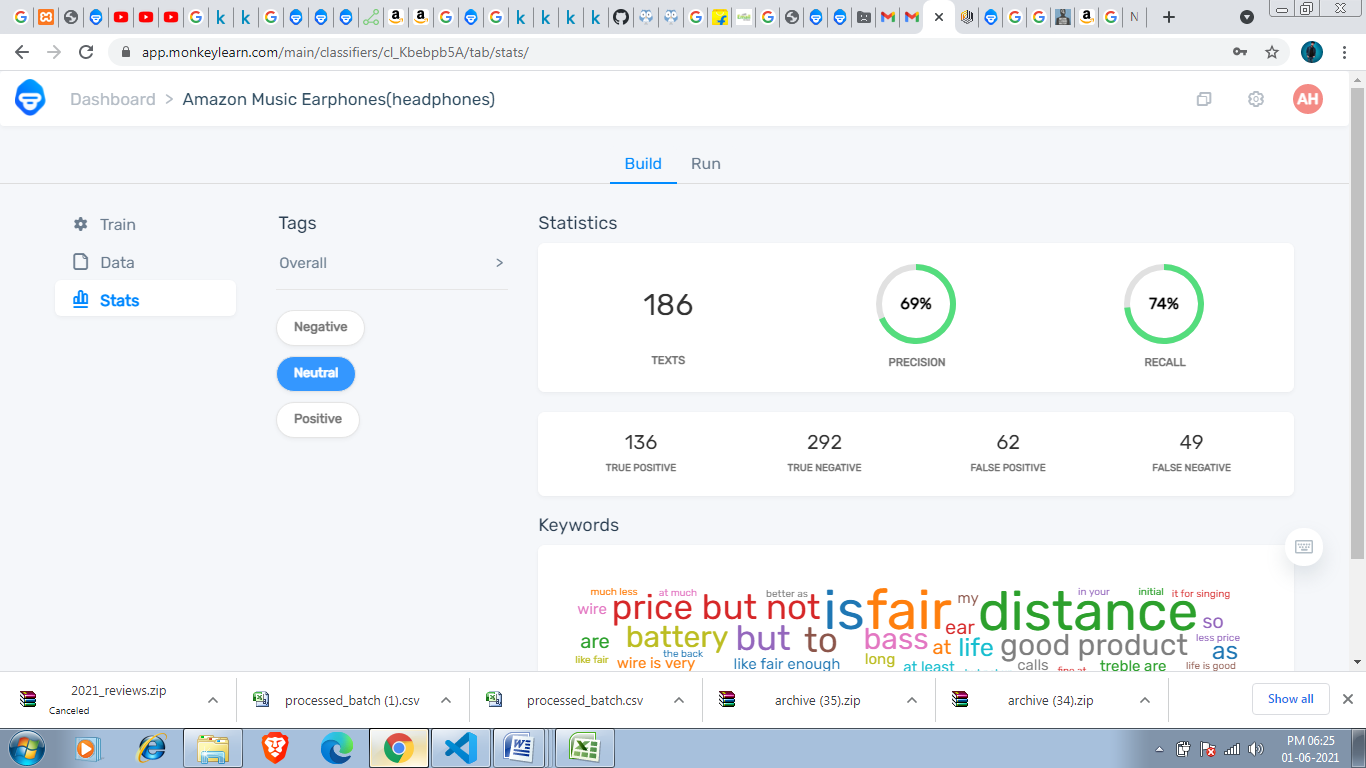
Overall Model accuracy



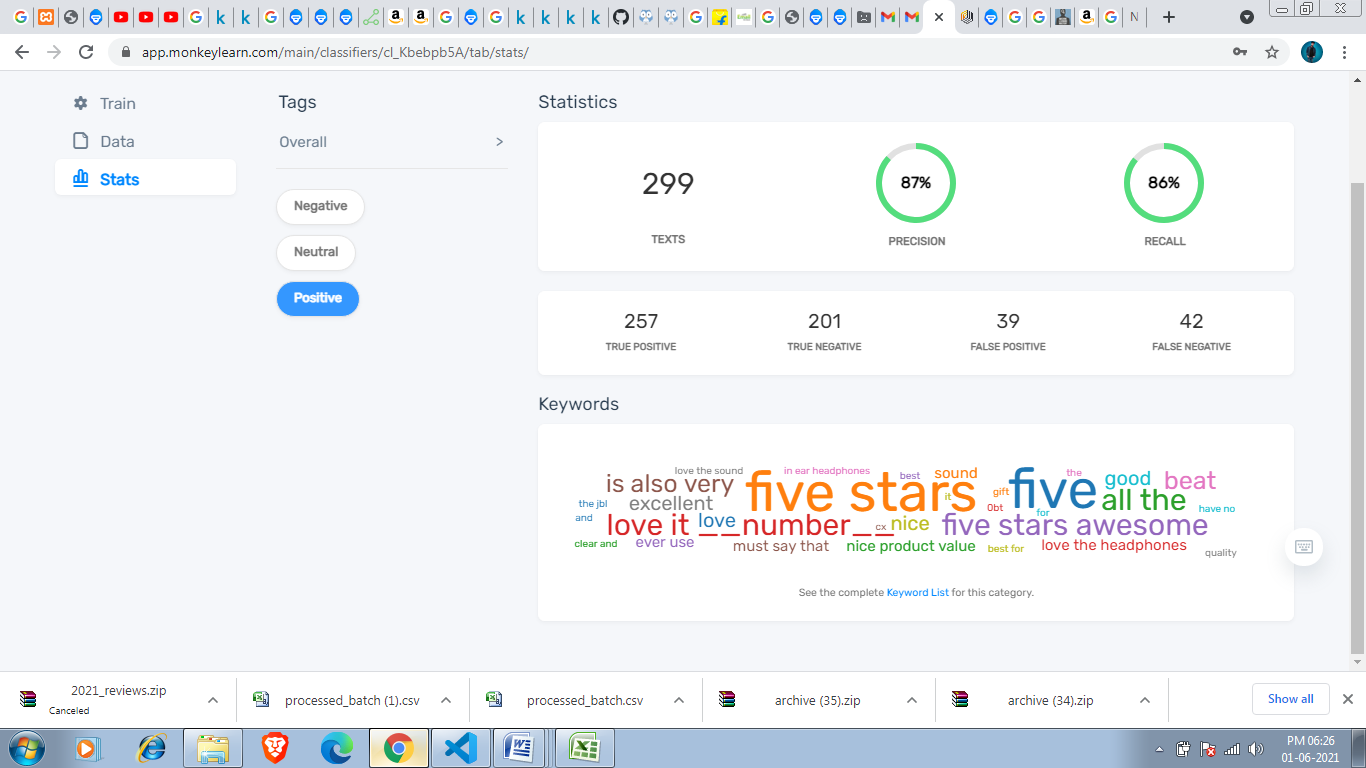
Negative tag precision:-



Neutral Tag Precision:-



Positive Tag precision:-



We can visualize this in Monkey learn Studio.