**Sinhgad Institute of Business**

**Administration and Research**

**MCA 1 YEAR (Sem 2)**



**Project Report**

**On**

**Comment Toxicity Scanner**

**By**

**Ojas Abhijit Hiwale**

**Roll No- 67**

**Project Guide: Snehal Dhane**

* 1. Introduction

The web is a harsh place and to go anonymous is very easy these days. People are more connected on social media. Thus, the amount of communications has also increased.

Thus, anything that is said over the internet can hurt the others feelings and the said person can easily get a away with it. This is not a straight solution to this as the internet is a free place and everyone has the right to say what they feel but they can be a way to try to identify miscreants that put things on social media to hurt others or to start conflicts.

This project is an attempt to identify such comments and flag them before they reach the masses.

2.2 Existing System and need for system

There are many systems that exist that can implement this current project many platforms like Facebook and YouTube specially do use such scanners or flaggers that can pinpoint hate comments. The problem with the pre-existing system is that people individually can also report comments that can be flagged down and in a lot of situations even if the comment is something that is useful it can be reported and then it has to go through some scrutiny and is then checked now instead of going through this system we can create a system wherein we can go ahead and resolve this without any human intervention at all decreasing the work load all together

2.3 Limitations of Existing System

Human intervention is needed in a lot of cases in a scenario that occurs where the machine won’t be able to handle it , human intervention is surely needed but the limitations to the pre-existing system is that again even if a comment that is useful or a statement that is useful or is meant to be in there if reported by a certain individual might be taken down instead of it being a better or a good comment so the pre-existing system has this limitation wherein it can be exploited in a way to ends to hurt others..

Proposed System

2.1 Problem Statement

The system or the model that is being used in this current project helps us understand as to how machine learning can help us reduce our workload. in this system we use a data set of pre-existing comments consisting of good comments as well as bad comments and they are rated according to 5 para metres through these five para metres the machine is then trained and then according to that it will go ahead and implement the system and flag the comments accordingly

We can go ahead and first identify such comments then rate them according to the toxicity and then go ahead and decide as to what we can do about this so this system can help us seclude a set of comments that are creating a problem over our entire system or our entire project and we can go ahead and decide as to what we can do about it. so this model can help us to filter out such comments and go ahead and decide if you want to flag them or how to go about them.

2.2 Summary Of capabilities

The system can Identify the difference between a good and appropriate comment and a toxic comment and filter them out. It is also very versatile as in we can make this better by making our database more in-depth and strong.

2.3 Objectives of the proposed System

The internet is a free place there are a lot of platforms that do not have any age restrictions so if anybody under the age of 18 can encounter a lot of wrong comment or inappropriate comments that can go ahead and hinder that individuals mindset this creates a problem for a lot of social media platforms because people will stop trusting them and because of a few individuals who are trying to create a problem so this model or the system helps us to pinpoint and filter out such comments so that we can go ahead and implement this in a lot of ways and this is a very versatile model as it does not take a lot of space and it can also be implemented with a standalone software in this system we have used the gradio software to implement this and it can also be implemented or tested publicly quite easily so this system in all is trying to filter out such comments so that we can go ahead and decide as to if you want them on a system or we want them out of the system

2.4 Functional Requirements

Python

Jupyter notebook

Gradio

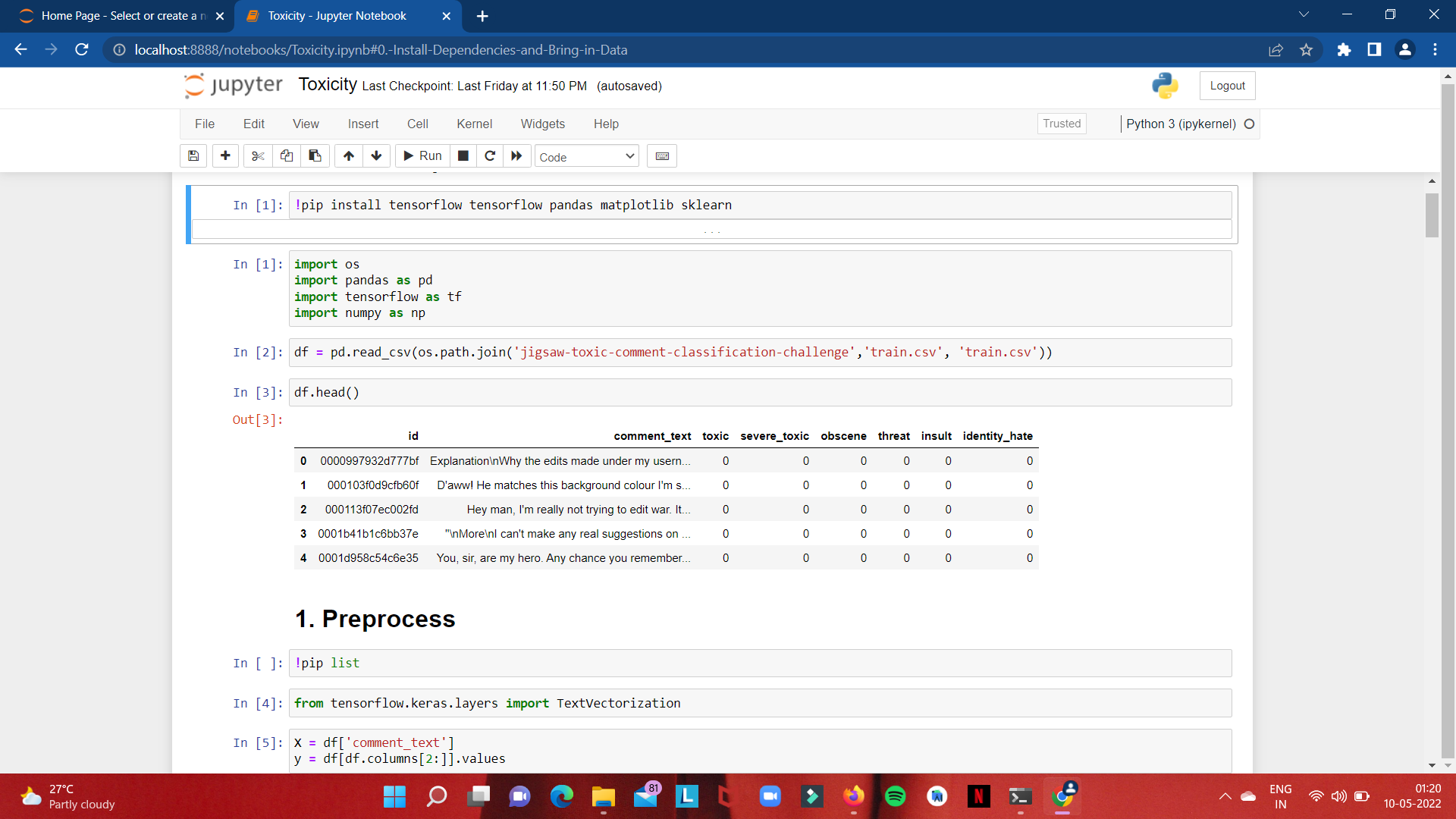
**Drawbacks and limitations**

1. The current system can do better with more data and training
2. This model can filter out from previous comments it was trained on if anything new comes up it won’t be able to handle.

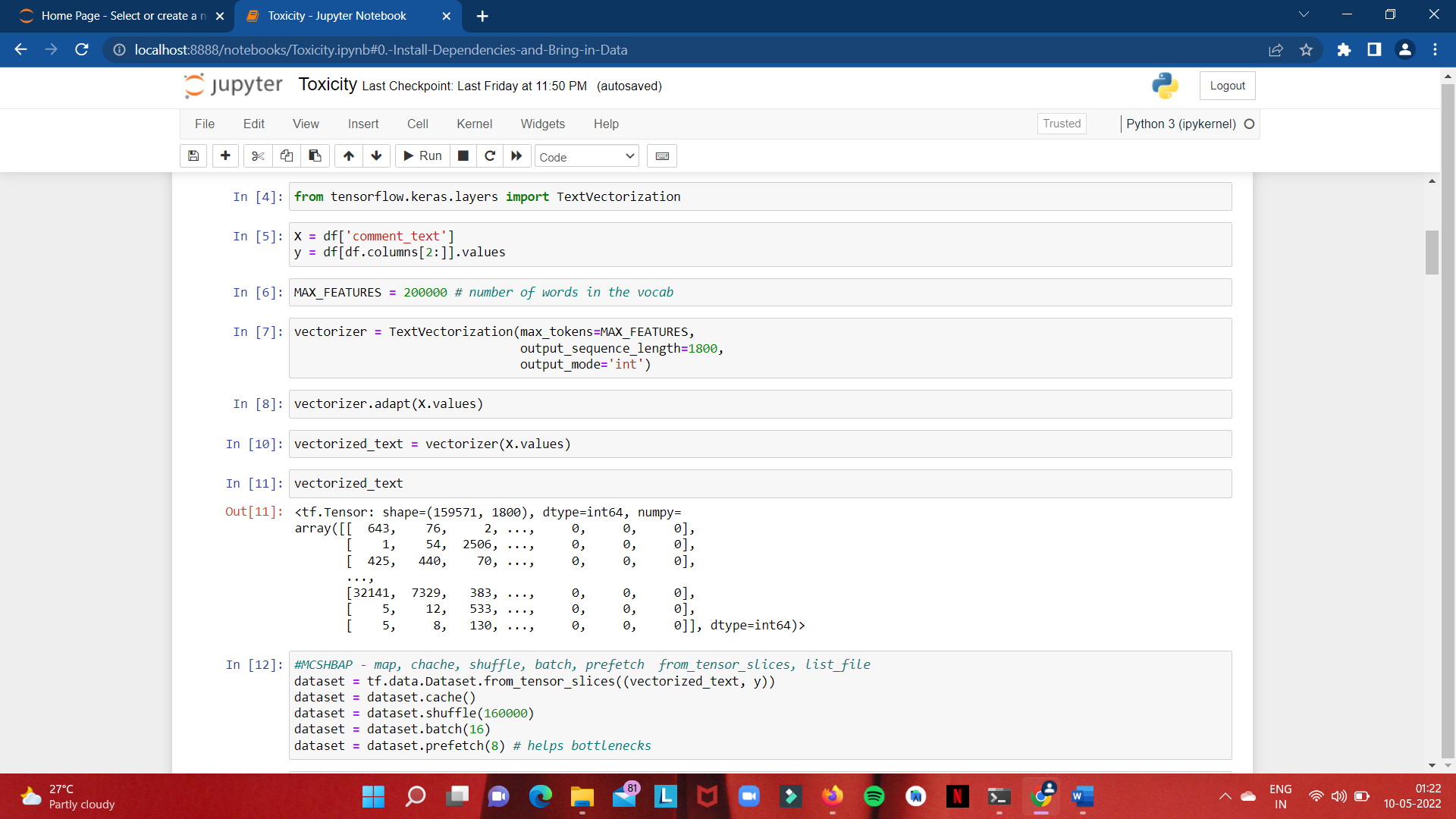
**Proposed Enhancements**

1. More datasets to train the model better.
2. Implements this in a real time environment.

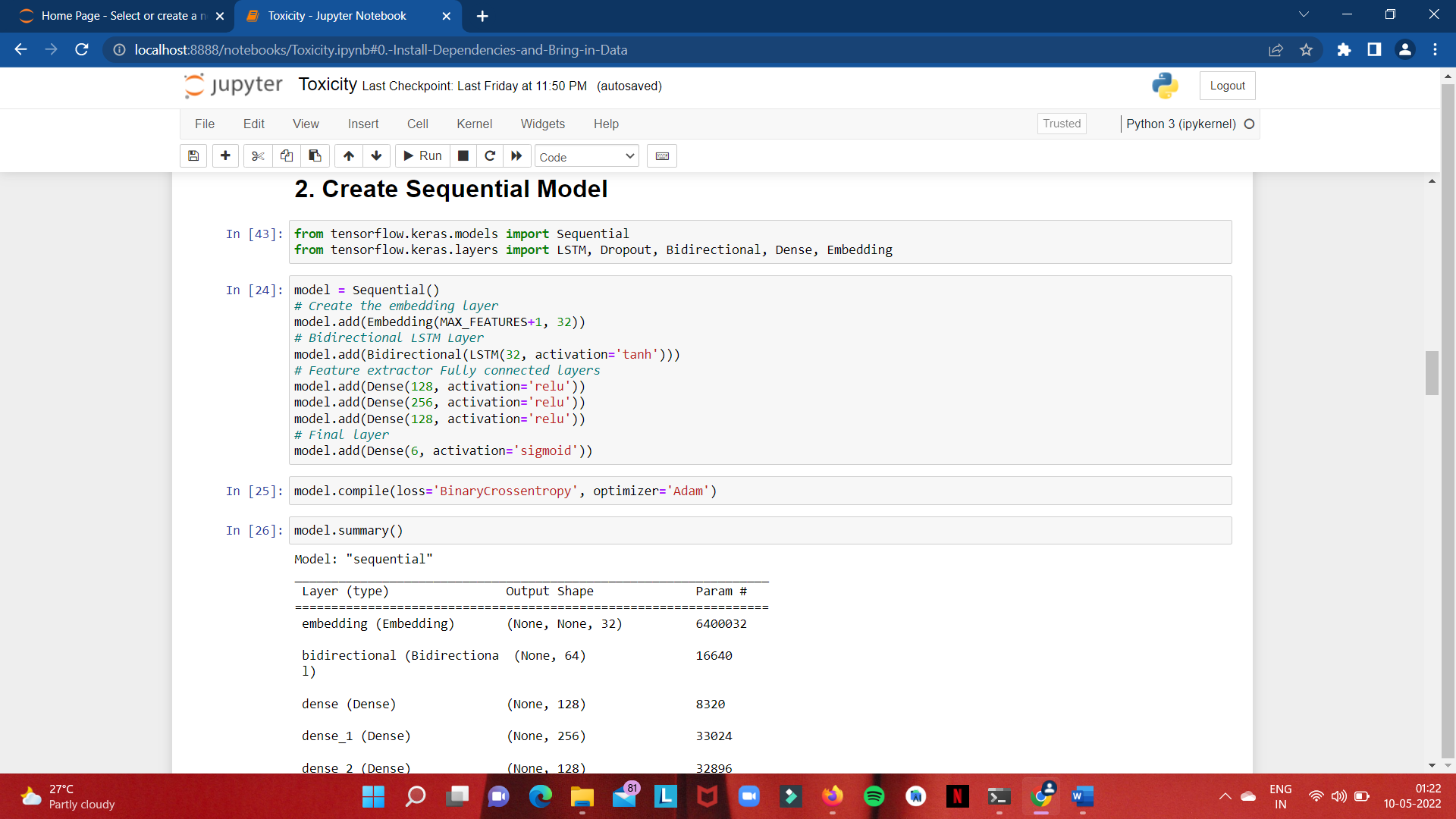
Importing the dependencies and dataset in the project



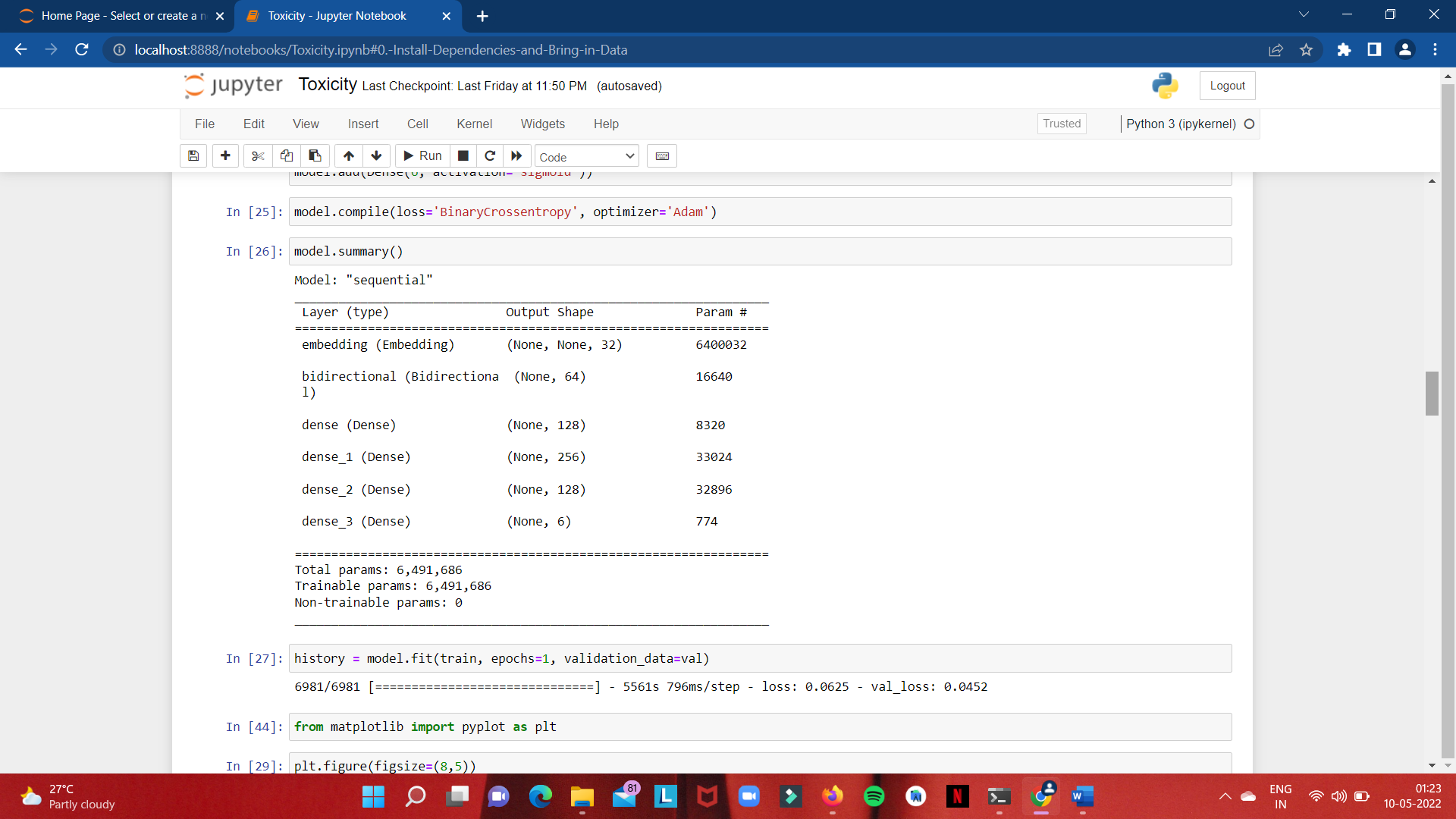
Vectorizing the data to fit the model

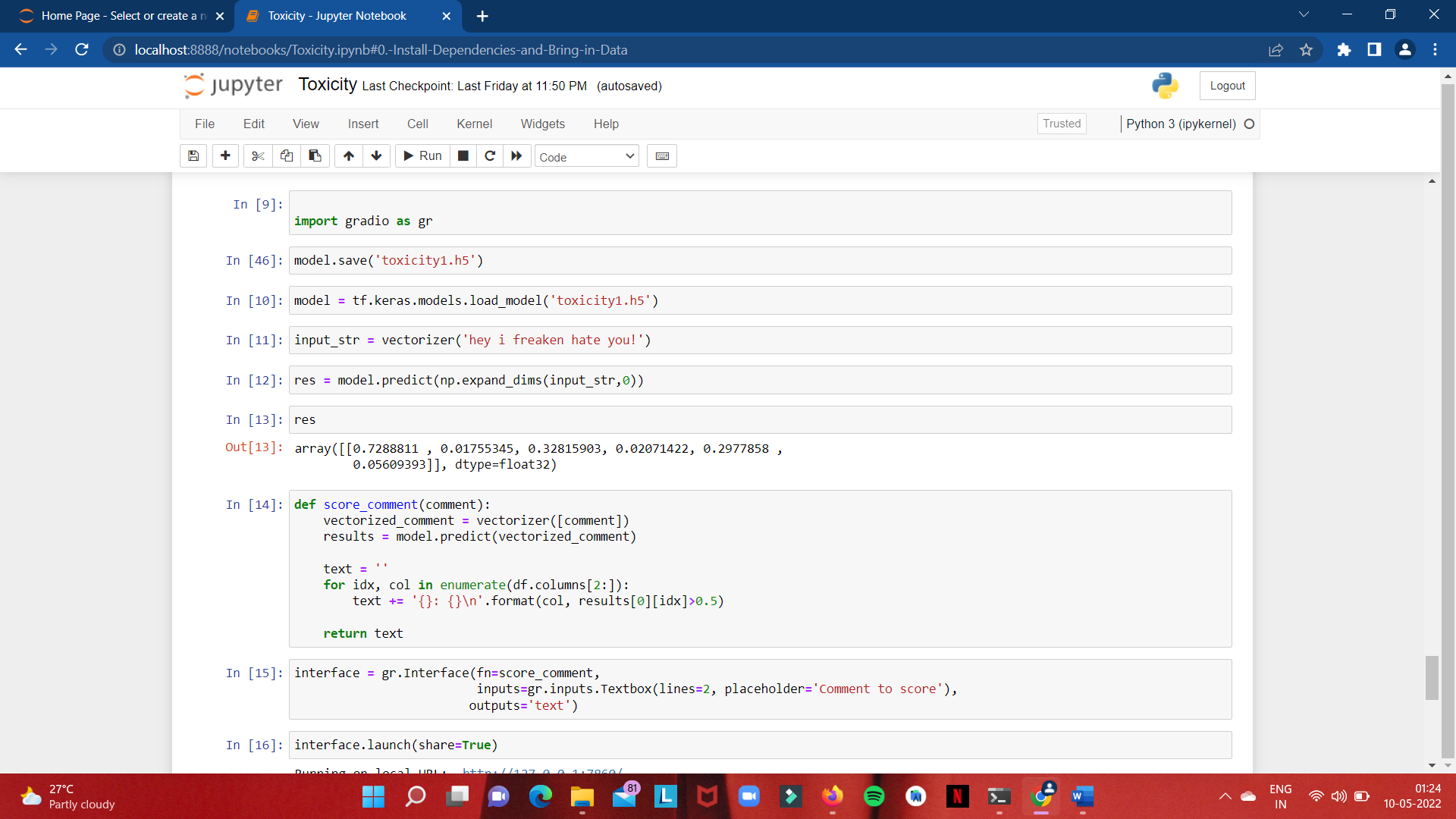


Creating the model

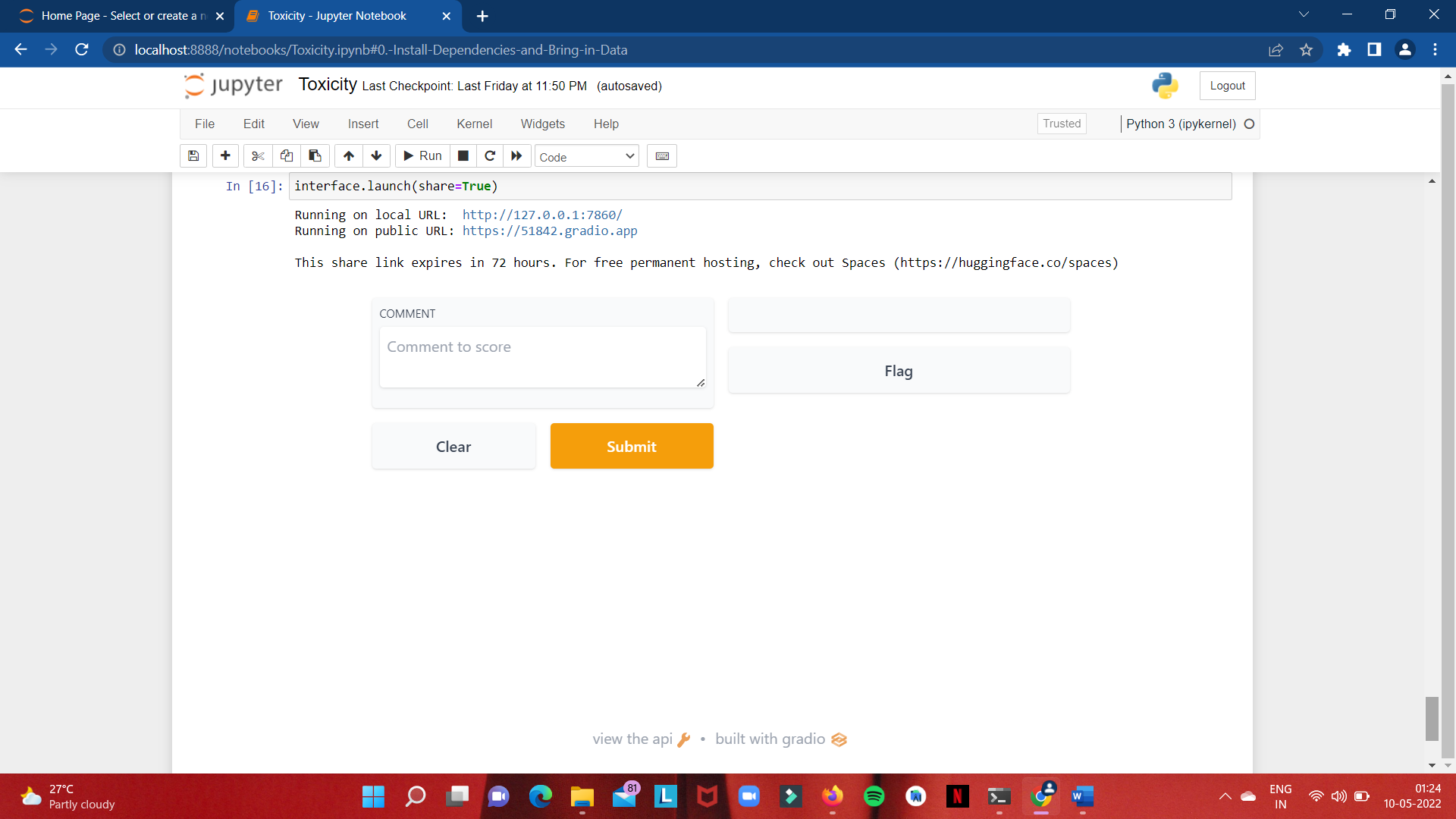


Model summary and model training

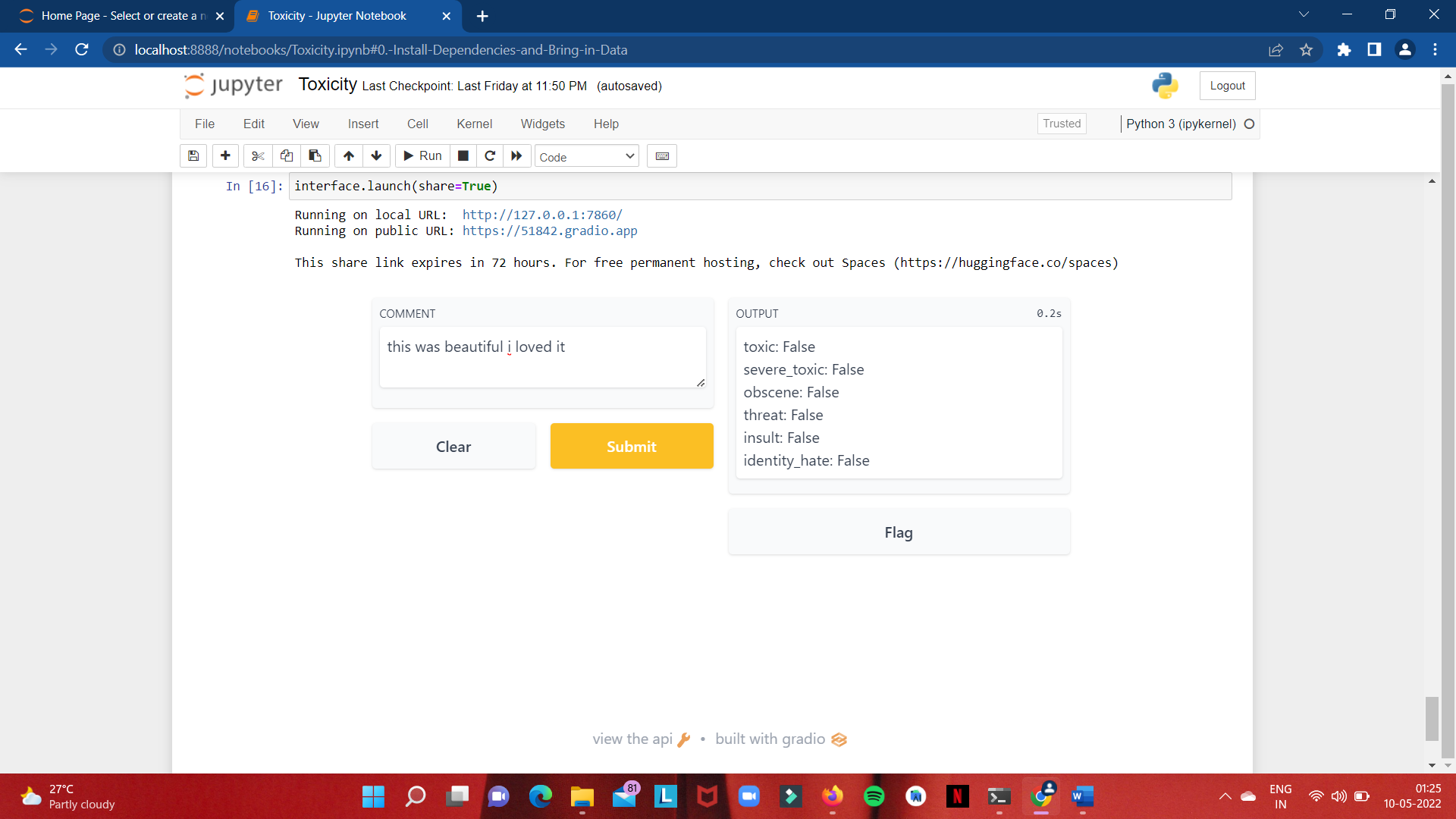




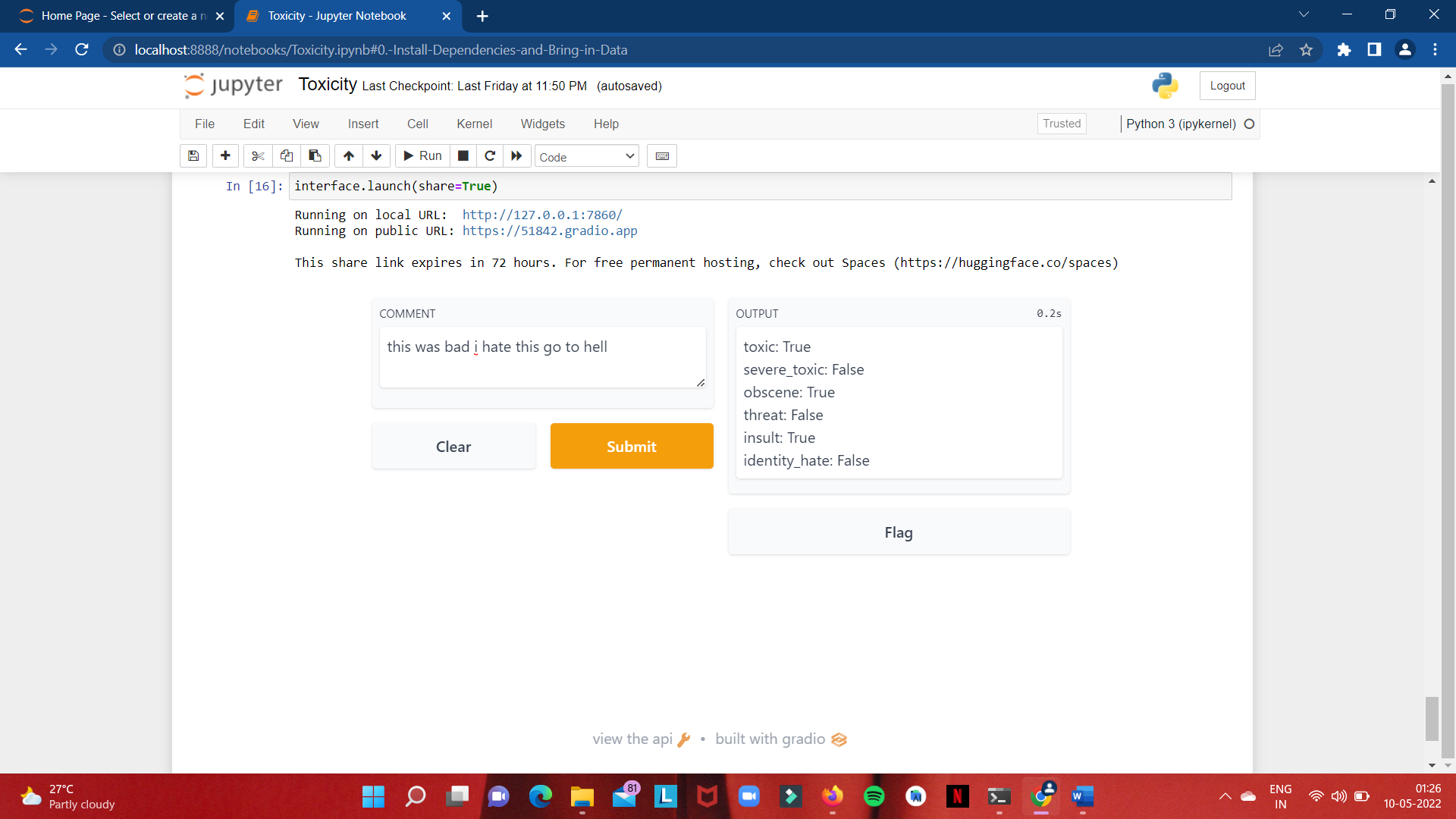
App interface



Good comment



Toxic comment



**Conclusion**

The main motive of this system is to filter out and identify comments that create a bad environment on the social media platforms a lot of times when we are trying to convey something through a text it does not actually go through or is misunderstood by a lot of people a lot of times.

On the other hand there are also comments that or intended and written in a way so that they can be hurtful so instead of that we can go ahead and use the system and try to filter out such problems and find a way around them

We can also go ahead and train this model to filter out and also solve the problems in conclusion this model can help us solve a real-time environment problem and at least make the workload on us quite less.

**Bibliography**

www.youtube.com

[www.github.com](http://www.github.com)

[www.stackoverflow.com](http://www.stackoverflow.com)

[www.google.com](http://www.google.com)