

# **Amazon Kindle Store Reviews Analysis**

**Using Deep Learning (DL)**

**Developed by: Arpita T M, Chandana Laxmi H, Nikita G Ghanate, Pooja Y U, Roja Devina.**

**SmartBridge - Remote Summer Internship Program**

## **1.INTRODUCTION**

Amazon Kindle Store is an e-book e-commerce store for all the book reading hobbyists. Online reviews are a category of product information created by the users based on personal handling experience. Online shopping websites endow with platforms for consumers to review products and carve up opinions. Sentiment analysis or opinion mining is nothing but classification of emotions in the reviews text into positive, negative and neutral. Amazon Kindle Store is an e-book e-commerce store for all the book reading hobbyists. Online reviews are a category of product information created by the users based on personal handling experience. Online shopping websites endow with platforms for consumers to review products and carve up opinions. Sentiment analysis or opinion mining is nothing but classification of emotions in the reviews text into positive, negative and neutral.

### **1.1 Overview**

With everything shifting online, reviews have started giving utmost importance to Sentiment Analysis. Honestly, it's their only gateway to thoroughly understanding their customer-base, including their expectations from the customer review. Social Media listening can help organisations from any domain understand the grievances and concerns of their customers – which eventually helps the organisations scale up their services. Sentiment Analysis helps brands tackle the exact problems or concerns of their customers.

Sentiment analysis gives an organisation the much-needed insights on their customers. Organisations can now adjust their marketing strategies depending on how the customers are responding to it. Sentiment Analysis also helps organisations measure the ROI of their marketing campaigns and improve their customer service. Since sentiment analysis gives the organisations a sneak peek into their customer's emotions, they can be aware of any crisis that's to come well in time – and manage it accordingly.

## SENTIMENT ANALYSIS



Discovering people opinions, emotions and feelings about  
a product or service

### 1.2 Purpose

Our aim for the project is to make use of Deep learning using python and extract necessary libraries from it for the sentiment analysis of the amazon kindle store reviews. The reviews dataset from the amazon kindle is to be pre-processed first by removing the punctuation marks and by applying the tokenization and attention masking step. This step is necessary to make the machine familiarise with the input text. The model is then trained under these texts and the model is built using the Convolutional neural networks. The main idea behind the project is to demonstrate the use of sentiment analysis for the customer feedback on each book from amazon kindle store.

## 2. LITERATURE SURVEY

Deep learning is a field of machine learning concerned with algorithms inspired by the structure of the brain called artificial neural networks. It has evolved hand-in-hand with the digital era, which has brought about an explosion of data in all forms and from every region of the world. This data, known simply as big data, is drawn from sources like social media, internet search engines, e-commerce platforms, and online cinemas, among others. This enormous amount of data is readily accessible and can be shared through fintech applications like cloud computing.

However, the data, which normally is unstructured, is so vast that it could take decades for humans to comprehend it and extract relevant information. Companies realize the incredible potential that can result from unraveling this wealth of information and are increasingly adapting to AI systems for automated support.

Each layer of its neural network builds on its previous layer with added data like a retailer, sender, user, social media event, credit score, IP address, and a host of other features that may take years to connect together if processed by a human being. Deep learning algorithms are trained to not just create patterns from all transactions, but also know when a pattern is signaling the need for a fraudulent investigation. The final layer relays a signal to an analyst who may freeze the user's account until all pending investigations are finalized.

Deep learning is used across all industries for a number of different tasks. Commercial apps that use image recognition, open-source platforms with consumer recommendation apps, and medical research tools that explore the possibility of reusing drugs for new ailments are a few of the examples of deep learning incorporation.

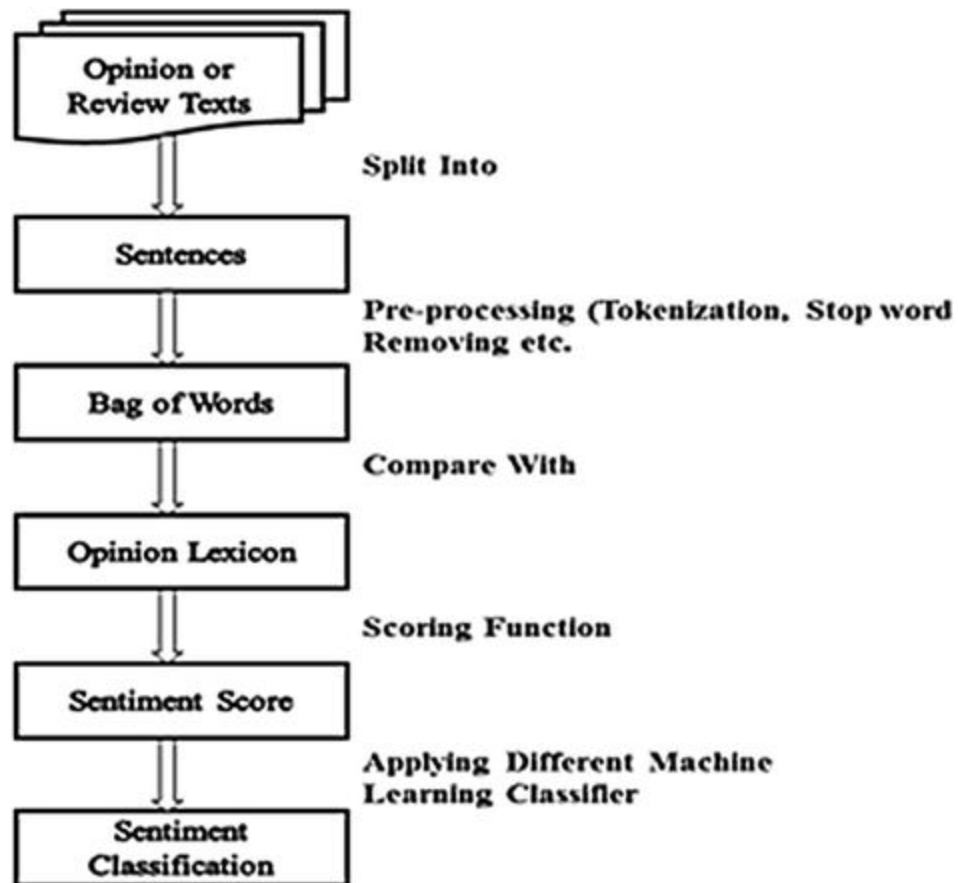
### 3. THEORETICAL ANALYSIS

While selecting the algorithm that gives an accurate prediction we went through a lot of algorithms which gave the results abruptly accurately and from them we selected only one algorithm for the prediction problem that is convolutional neural network.

#### 3.1 Block Diagram



**Fig: Outline of Classification in Machine Learning**



**Fig: Block diagram of REVIEW SENTIMENT ANALYSIS**

### **3.2 Software Designing**

- Jupyter Notebook Environment
- Spyder Ide
- Machine Learning Algorithms
- Python (pandas, numpy,keras,Tensorflow,CNN layers)
- HTML
- Flask

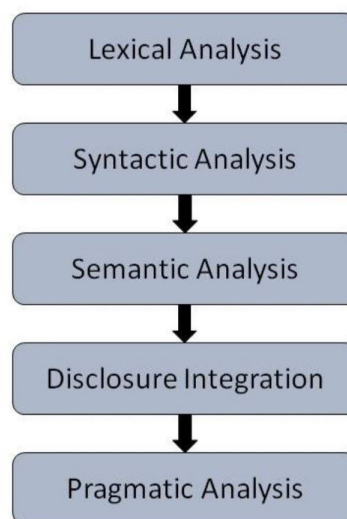
We developed this loan status prediction by using the Python language which is an interpreted and high level programming language and using the Machine Learning algorithms. For coding we used the Jupyter Notebook environment of the Anaconda distributions and the Spyder, it is an integrated scientific programming in the python language. For creating an user interface for the prediction we used the Flask. It is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions, and a scripting language to create a web page is HTML by creating the templates to use in the functions of the Flask and HTML.

## 4. EXPERIMENTAL INVESTIGATION

In this paper, the dataset we used for the purpose of doing sentiment analysis is derived from <https://www.kaggle.com/bharadwaj6/kindle-reviews>. In this we have rows and columns. It contains the books names and the respective review for all the books and also the other information related to the book and also the user. To review the comments we make use of the Natural Language Processing (NLP) algorithm it is a Deep learning algorithm which can process the sentences and remove the all punctuations and other string processing is done by this algorithm.

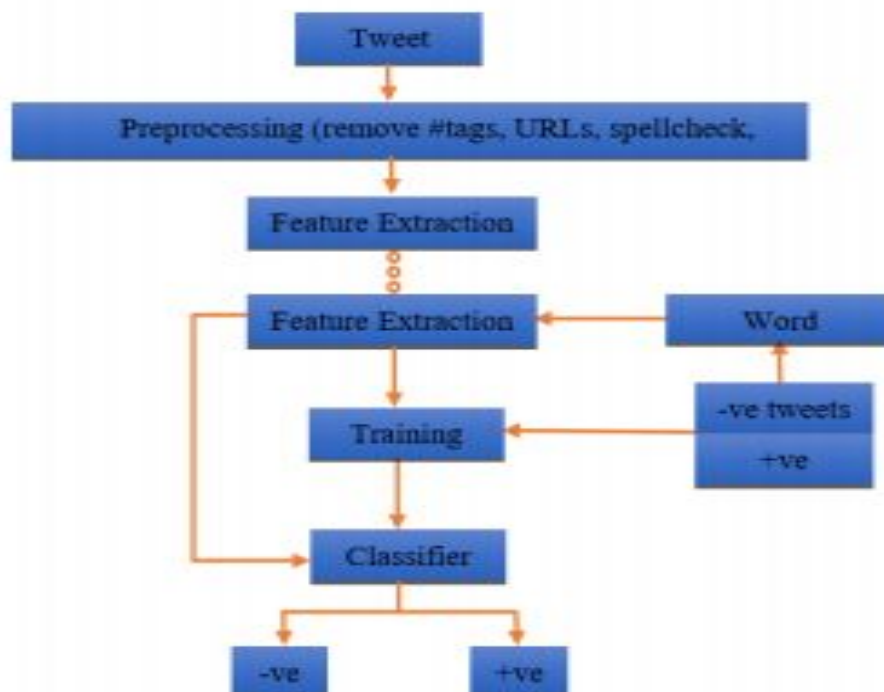
### Steps in NLP:

There are five general steps.



- **Lexical Analysis** – It involves identifying and analyzing the structure of words. Lexicon of a language means the collection of words and phrases in a language. Lexical analysis is dividing the whole chunk of text into paragraphs, sentences, and words.
- **Syntactic Analysis (Parsing)** – It involves analysis of words in the sentence for grammar and arranging words in a manner that shows the relationship among the words. The sentence such as “The school goes to boy” is rejected by English syntactic analyzers.
- **Semantic Analysis** – It draws the exact meaning or the dictionary meaning from the text. The text is checked for meaningfulness. It is done by mapping syntactic structures and objects in the task domain. The semantic analyzer disregards sentences such as “hot ice-cream”.
- **Discourse Integration** – The meaning of any sentence depends upon the meaning of the sentence just before it. In addition, it also brings about the meaning of immediately succeeding sentences.
- **Pragmatic Analysis** – During this, what was said is re-interpreted on what it actually meant. It involves deriving those aspects of language which require real world knowledge.

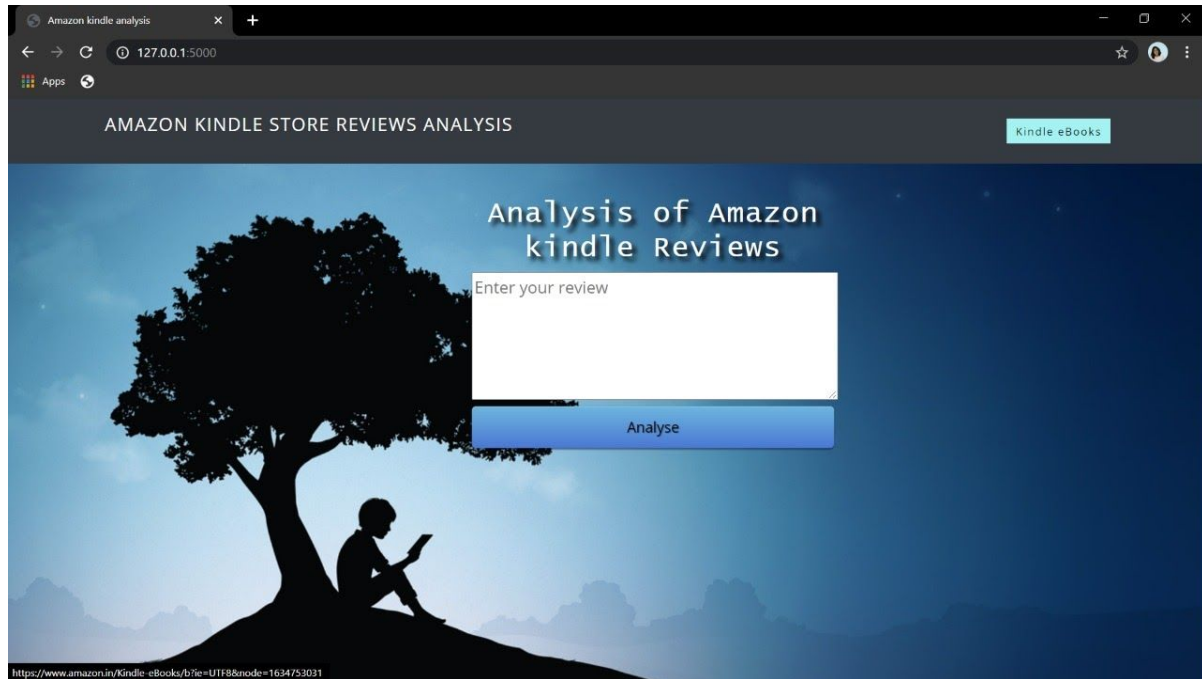
## 5.FLOWCHART



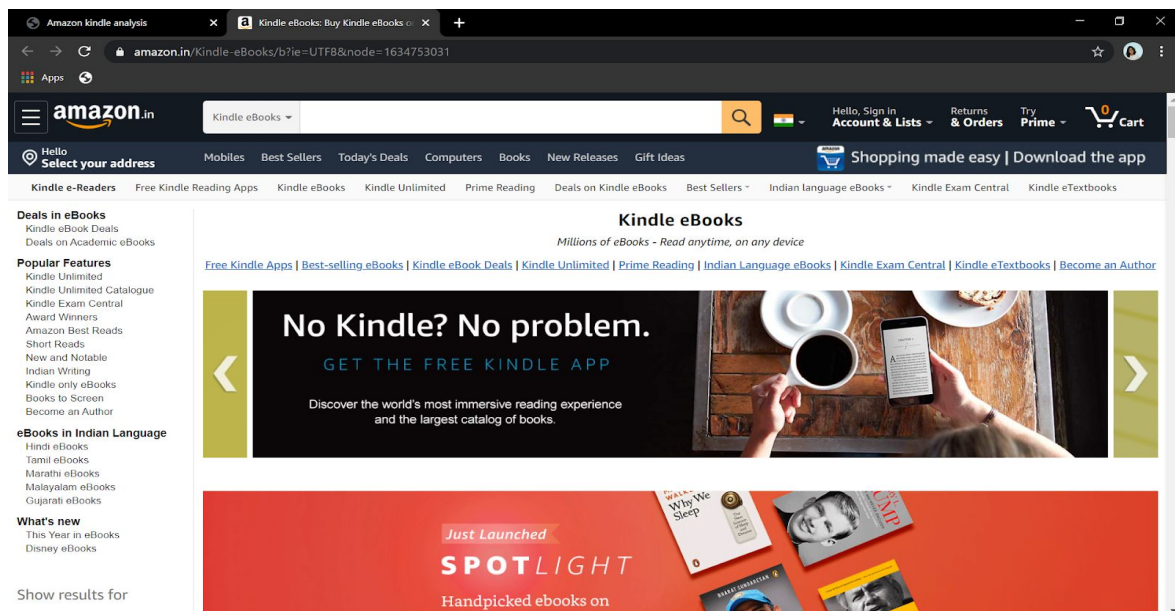
## 6. RESULT

The following are the snapshots of the resulting model proposed by the NLP to do the sentiment analysis on the reviews directed from the amazon kindle book store.

**Fig1: Home page**

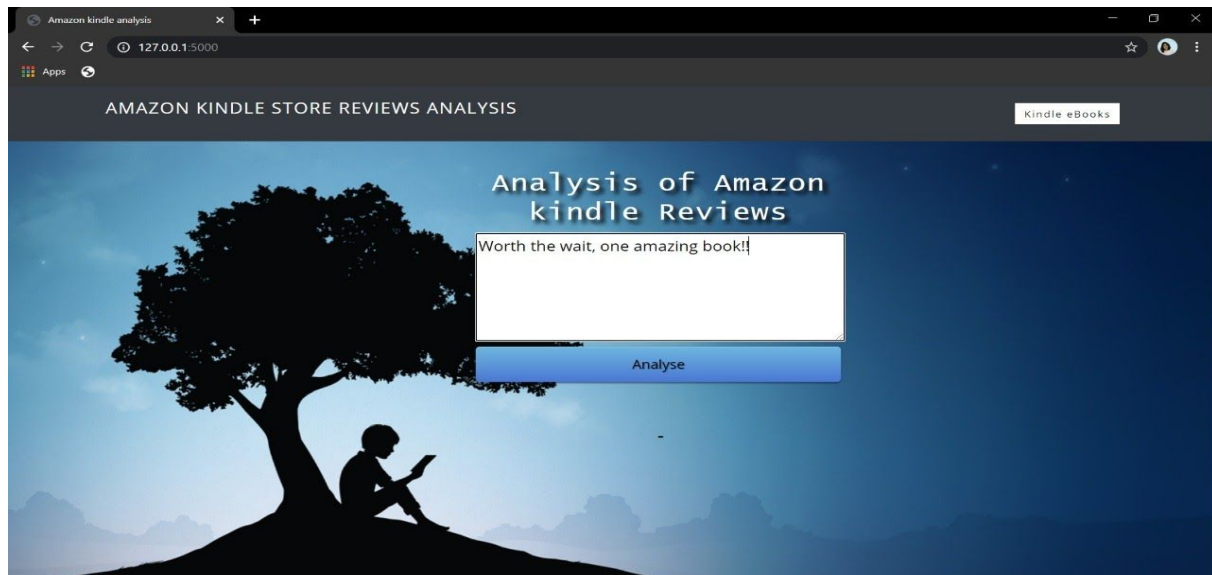


This is the home page. The user will enter the command which he wants to analyze and if user click on the Kindle eBooks button then it redirect to the screen shown in below fig





**Fig2:**



The user enters the review in the review box and press the analyse button to know the sentiment of the review.

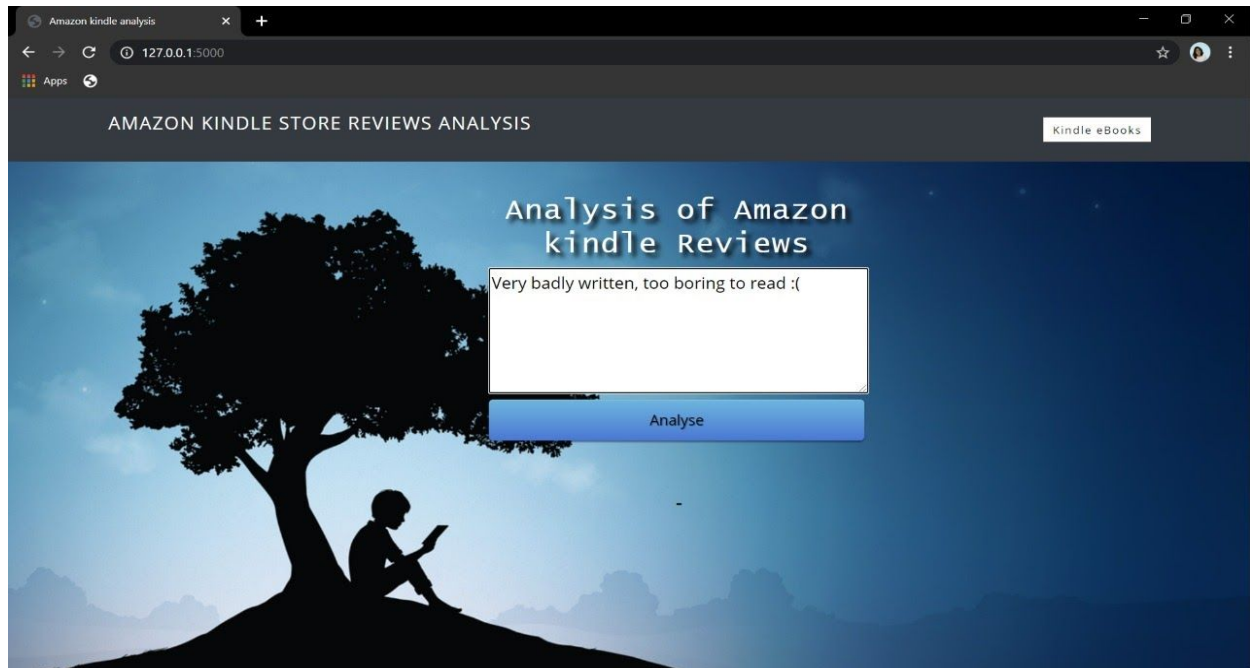
**Fig3:**



The user gets the output as the **Positive review**.



**Fig4:**



The user entered the another review in the review box and clicked on analyse button  
He will be redirected for the output page I.e fig 5.

**Fig5:**



The user gets the output as a **Negative review**.

## **7. ADVANTAGES AND DISADVANTAGES**

### **ADVANTAGES:**

- Happy customers are more likely to be receptive to upselling. With sentiment analysis, you can easily identify your happiest customers. This helps you recognise chatters who might be receptive to spending more, as well as avoiding upsetting disgruntled customers with any unwelcome sales pitches.
- Users can ask comments about any book and get the direct response within seconds.
- The NLP system provides answers to sentences or questions in natural language.
- This system often has an exact answer to the questions. No unnecessary or unwanted information.
- The accuracy of the answers increases with the amount of relevant information provided in the question.
- Rapid prototyping and testing.
- Robust system in the NLP will always produce outputs, regardless of the image entered.
- It produces output in relation to the data entered.

### **DISADVANTAGES:**

- Long time to implement and test a prototype.
- Manual system creation time-consuming and expensive creation process.

## **8. APPLICATION**

- Text-based application searching that the comment is positive or negative. extracting the emotion(positive or negative) of the given sentence by the users.
- Dialogue based application answering system.

## 9. CONCLUSION

It is completely impossible to use only raw text as input for making predictions. Hence, we saw that pre-processing step played a major role in the complete process of NLP. To get better results, accuracy and make the machine take all the text as tokens, pre-processing of data is to be done carefully looking at the type of contents present in it.

The most important thing is to be able to extract the relevant features from the given source of data. This kind of data can often come as a good complementary source in order to extract more learning features and increase the predictive power of the models. And the user is able to predict that the given comment is positive or negative.

## 10. BIBLIOGRAPHY

1. Gonzalez-de Santos P, et al. Fleets of robots for environmentally-safe pest control in agriculture. *Precis. Agric.* 2017;18:574–614. doi: 10.1007/s11119-016-9476-3.
2. Fernández-Quintanilla C, et al. Is the current state of the art of weed monitoring suitable for site-specific weed management in arable crops? *Weed Res.* 2018;58:259–272. doi: 10.1111/wre.12307.
3. Commonwealth of Australia. Agricultural competitiveness white paper. ISBN: 978-1-925237-73-3 (2015).
4. Slaughter DC, Giles DK, Downey D. Autonomous robotic weed control systems: A review. *Comput. Electron. Agric.* 2008;61:63–78. doi: 10.1016/j.compag.2007.05.008.
5. Shaner DL, Beckie HJ. The future for weed control and technology. *Pest Manag. Sci.* 2014;70:1329–1339. doi: 10.1002/ps.3706.
6. Bakhshipour A, Jafari A. Evaluation of support vector machines and artificial neural networks in weed detection using shape features. *Comput. Electron. Agric.* 2018;145:153–160. doi: 10.1016/j.compag.2017.12.032. [[CrossRef](#)] [[Google Scholar](#)]
7. dos Santos Ferreira A, Freitas DM, da Silva GG, Pistori H, Folhes MT. Weed detection in soybean crops using ConvNets. *Comput. Electron. Agric.* 2017;143:314–324. doi: 10.1016/j.compag.2017.10.027. [[CrossRef](#)] [[Google Scholar](#)]
8. Dyrmann M, Jørgensen RN, Midtiby HS. RoboWeedSupport - Detection of weed

locations in leaf occluded cereal crops using a fully convolutional neural network. *Adv. Animal Biosci.* 2017;8:842–847. doi: 10.1017/S2040470017000206. [[CrossRef](#)] [[Google Scholar](#)]

9. Wu, S. G. *et al.* A leaf recognition algorithm for plant classification using a probabilistic neural network. In *Proceedings of the 2007 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 11–16 (Cairo, Egypt, 2007).

10. Kumar, N. *et al.* Leafsnap: A computer vision system for automatic plant species identification. In *Proceedings of the 2012 European Conference on Computer Vision (ECCV)*, 502–516 (Berlin, Heidelberg, 2012).

## APPENDIX

### HTML:

#### Base.html:

```
<html lang="en">
```

```
<head>
```

```
  <meta charset="UTF-8">
```

```
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
```

```
  <meta http-equiv="X-UA-Compatible" content="ie=edge">
```

```
  <title>Amazon kindle store review analysis</title>
```

```
    <link    href="https://cdn.bootcss.com/bootstrap/4.0.0/css/bootstrap.min.css"
rel="stylesheet">
```

```
  <script src="https://cdn.bootcss.com/popper.js/1.12.9/umd/popper.min.js"></script>
```

```
  <script src="https://cdn.bootcss.com/jquery/3.3.1/jquery.min.js"></script>
```

```
  <script src="https://cdn.bootcss.com/bootstrap/4.0.0/js/bootstrap.min.js"></script>
```

```
  <link href="{{ url_for('static', filename='css/main.css') }}" rel="stylesheet">
```

```
<style>

body {

    background-image: url("https://images3.alphacoders.com/105/thumb-1920-105099.jpg");

    background-color:#B3C99E

    background-size:cover;

    background-repeat: no-repeat;

    background-position:center bottom;

}

</style>

</head>

<body>

<nav class="navbar navbar-dark bg-dark">

    <div class="container">

        <a class="navbar-brand" href="#">AMAZON KINDLE STORE REVIEW
        ANALYSIS</a>

        <a href="https://software.farm.bot/docs/weed-detection" class="w3-btn w3-black"
        type="button">Help</a>

    </div>

</nav>

<div class="container">
```

```
<div id="content" style="margin-top:2em">{% block content %}{% endblock %}</div>
```

```
</div>
```

```
</body>
```

```
<footer>
```

```
<script src="{{ url_for('static', filename='js/main.js') }}" type="text/javascript"></script>
```

```
</footer>
```

```
</html>
```

### **Index.html:**

```
{% extends "base.html" %} {% block content %}
```

```
<h1>Sentiment analysis</h1>
```

```
<div>
```

```
<br>
```

```
<form id="upload-file" method="post" enctype="multipart/form-data">
```

```
<label style="color:black" for="imageUpload" class="upload-label">
```

```
    Choose...
```

```
</label>
```

```
<input type="file" name="file" id="imageUpload" accept=".png, .jpg, .jpeg">
```

```
</form>
```

```
<div class="image-section" style="display:none;">
```

```
<div class="img-preview">
```

```

        <div id="imagePreview">

        </div>

    </div>

    <div>

    <button type="button" class="btn btn-primary btn-lg" id="btn-predict">Predict!</button>

    </div>

</div>

<div class="loader" style="display:none;"></div>

<h2 id="result" style="background-color:white;color:black;font-size:16px;">

    <span></span>

</h2>

</div>

{% endblock %}

```

## APP.PY:

```

import numpy as np
from flask import Flask, request, render_template
from joblib import load
import joblib
from tensorflow.keras.models import load_model
from sklearn.feature_extraction.text import CountVectorizer

```



```

app = Flask(__name__)
model=load_model("amazo.h5")
@app.route('/')
def home():
    return render_template('index.html')
@app.route('/y_predict',methods=['POST'])
def y_predict():
    """
    For rendering results on HTML GUI
    """
    d = request.form['Sentence']
    print(d)

loaded=CountVectorizer(decode_error='replace',vocabul
ary=joblib.load('amazo.save'))
d=d.split("delimiter")
result=model.predict(loaded.transform(d))
prediction=result>0.5
if prediction[0] == False:
    output="Positive review"
elif prediction[0] == True:
    output="Negative review"
    return render_template('index.html',
prediction_text='{}'.format(output))
if __name__ == "__main__":
    app.run(debug=True)

```