

Text sentiment classifier proposal

Genuine challenge S4-AI | Luuk de Kinderen

“Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean commodo ligula eget dolor.”



| | |
|--|-----------|
| Project goal | 3 |
| Main question | 3 |
| Project plan | 3 |
| Context understanding | 4 |
| Review | 4 |
| User Review | 4 |
| Bought review | 4 |
| Positive or negative sentiment | 5 |
| Negative | 5 |
| Positive | 5 |
| Potential Impact assessment | 6 |
| Impact on society | 6 |
| Stakeholders | 7 |
| Privacy | 9 |
| Transparency | 10 |
| Inclusivity | 11 |
| Sustainability | 12 |
| Hateful and criminal actors | 13 |
| Human values | 14 |
| Future | 15 |
| Potential Impact assessment conclusion | 16 |
| Datasets | 17 |
| Movies | 17 |
| Hotels | 17 |
| Amazon products | 17 |
| Modelling | 18 |
| Linear SVC | 18 |
| Naive bayes | 18 |
| Evaluation | 19 |
| Domain knowledge verification | 19 |
| Results | 19 |
| Deployment | 19 |
| Literature list | 20 |

Project goal

Several sites allow you to post a comment or opinion about a product, place, or event. Think of Facebook, Instagram, Youtube, or Reddit. However, on most of these sites, you can not see quickly the sentiment of these comments. We want to know if these comments are written with a positive or negative mindset.

For many people or companies, it is useful to know what percentage of the reactions are positive or negative. You can't always tell this by looking at a Like system.

Main question

The main question you should be able to answer with the help of this project is; *What is the most effective way to classify text sentiment?*

Project plan

To achieve this goal, I am going to use reviews. The advantage of reviews is that they are usually already labeled. For example, you can directly translate a star rating to positive or negative. And many review datasets are already labeled as positive and negative.

Context understanding

Review

A review is an evaluation of a publication, service, or company such as a movie, video game, musical composition, book; a piece of hardware like a car, home appliance, or computer; or an event or performance, such as a live music concert, play, musical theater show, dance show, or art exhibition.

(Wikipedia, 2021)

User Review

A user review refers to a review written by a user or consumer of a product or a service based on her experience as a user of the reviewed product. Popular sources for consumer reviews are e-commerce sites like Amazon.com, Zappos or lately in the Yoga field for schools such as Banjaara Yoga and Ayurveda, and social media sites like TripAdvisor and Yelp. E-commerce sites often have consumer reviews for products and sellers separately. Usually, consumer reviews are in the form of several lines of texts accompanied by a numerical rating. This text is meant to aid in shopping decision of a prospective buyer. A consumer review of a product usually comments on how well the product measures up to expectations based on the specifications provided by the manufacturer or seller. It talks about performance, reliability, quality defects, if any, and value for money. Consumer review, also called 'word of mouth' and 'user generated content' differs from 'marketer generated content' in its evaluation from consumer or user point of view. Often it includes comparative evaluations against competing products. Observations are factual as well as subjective in nature. Consumer review of sellers usually comment on service experienced, and dependability or trustworthiness of the seller. Usually, it comments on factors such as timeliness of delivery, packaging, and correctness of delivered items, shipping charges, return services against promises made, and so on.

Consumer reviews online have become a major factor in business reputation and brand image due to the popularity of TripAdvisor, Yelp, and online review websites. A negative review can damage the reputation of a business and this has created a new industry of reputation management where companies attempt to remove or hide bad reviews so that more favourable content is found when potential customers do research.

(Wikipedia, 2021)

Bought review

A bought review is the system where the creator (usually a company) of a new product pays a reviewer to review his new product. Primarily used in the car, movie, and game industry this system creates a kind of undercover advertising. Bought reviews tend to be biased due to the informative value of reviews. In some cases, a bought review may be independent, if the person that is hired to do the review has a strong reputation for independence and integrity. Even if a "bought review" from a respected critic is actually independent, the perception of potential bias will remain, due to the financial relationship between the

company and the critic.
(Wikipedia, 2021)

Positive or negative sentiment

What exactly makes text positive or negative? There are very many components that determine the sentiment of text. But by looking at certain keywords, you can most easily guess the sentiment of the text. Think about words like "good" or "bad." These words by themselves are literal translations of sentiment.

Negative

You can often use other components to guess the sentiment of the text. Negative text, for example, is more often written in the past tense. When people are negative about something they have more often than not already put it behind them. So it is no longer a problem of the present. For example; "This product was very bad" is written in the past tense. While; "This product is very good" is written in present tense.

Negative reviews also use contradictions more often. Signal words such as "but" and "or" are much more common in negative text than in positive text. For example; "It was okay, but it could have been much better." Is clearly a negative text.

Positive

It is noticeable that there are many more keywords for positive text than negative text. Think "lovely", "great", "excellent" etc. These are all words that are a direct translation of a positive atmosphere/setting.

In addition, you can recognize positive text by one more thing. When someone is positive about something he/she also wants to tell more about it. So when people express themselves positively about something or someone they use more words on average. In addition, when someone expresses themselves positively, more positive things are mentioned. The keyword "and" is also more often found in positive text.

Potential Impact assessment

Impact on society

- *What is the challenge at hand? What problem (what 'pain') does this technology want to solve*
Using the technology a user can predict the sentiment of a piece of text without even reading the text.
- *Can you motivate why you are sure that this technology is solving the right problem?*
It can be very useful for different people, to know within a few seconds whether the text is positive or negative. If you quickly want to know the average opinion about something or someone. Then this tool is perfect to use.
- *In what way is this technology contributing to a world you want to live in?*
In this world, we want to move fast. Getting results as fast as possible is important to compete with others. This technology helps the world to move even faster. And that alone really excites me.
- *Now that you have thought hard about the impact of this technology on society, what improvements would you like to make? List them below.*
 1. Combine this technology with a web scraper. That way manual input of multiple texts becomes obsolete. And getting results will be even faster.
 2. Adding a percentage calculator while using more inputs. If you have multiple inputs of different pieces of text. Knowing what percentage of the texts is classified as positive or negative will be really interesting to add.
 3. Showing the classification certainty while making conclusions. This certainty can be interesting for the users.

Stakeholders

- *What are the main users/target groups/stakeholders for this technology?*

- **Companies that produce/deliver products/services**

By having a clear overview of the average overall experience of a product/service early on, companies can adapt quickly if they know in time that the average opinion of their own product/service has changed. In addition, it can be helpful to know the average opinion about similar products/services of competing companies. In fact, a company can compete better when they know the overall opinion of their competitor.

This stakeholder is a large group, but a very important one. If this technology is created, it will have a big impact on the market they are in. So this stakeholder is definitely one to take into account.

- **Companies that sell products/services**

If you sell multiple products/services, it is very helpful to know which of these are more popular with customers than others. As a company, you can then better advertise these products/services.

- **Consumers of products/services**

Of course, if using this technology changes the development and promotion of products/services, it also affects consumers. The choice between products/services will change.

- **Pollsters**

This technology will be a revolution for pollsters. Not taking into account this stakeholder is unavoidable. It will be a big change because much of their work will be automated and this will simplify their work. However, this does not mean that it will be positive for everyone, as it may mean that jobs will disappear.

- **News channels**

When a news channel wants to convey an insight of opinions to the viewer. The help of pollsters is often enlisted. With the help of this technology, this will become unnecessary in some cases. Channels will be able to come up with results more quickly themselves.

- **Politicians**

People's opinion about an idea or political proposal is very important for politicians to know. If the majority of all people are negative about a proposal, it can have dramatic consequences for a politician and his/her entire party. Of course, this also works the other way around. If a politician or party comes up with an idea that is well received by everyone, this may mean that a party will grow in voters.

When the use of this technology becomes mainstream. It would mean that politicians would handle the opinion of voters differently. Everyone can see

the average opinion of voters and this can also be used as an argument in a debate.

This stakeholder is certainly important to take into account because it has a big impact on their way of working. And the work of politicians, of course, has an effect on everyone.

- *Now that you have thought hard about all stakeholders, what improvements would you like to make? List them below.*
 1. Adding a warning that the classification is a prediction would be a good addition. By adding this warning you prevent users from using the results as a fact.
 2. It might be a good addition that when a prediction is published, it should include a source. Such as: "According to the opinion forecaster". This will prevent the technology from being used as a fact.

Privacy

- *Does this technology register personal data? If yes, what personal data?*
no, this technology only uses personal data to come to a prediction. This data is not stored. Whether the data is actually personal is also a matter of debate. It only uses a piece of text and no name is attached to it.
- *Do you think this technology invades someones privacy? If yes, in what way?*
When this technology is used by someone. It may be that another person's opinion is used to make a conclusion. This could be interpreted as a privacy violation. But this was also possible without this technology. The comments of others are always visible and conclusions can be drawn from them. The only thing that has changed is that this technology makes this easier and faster.
- *Do you think this technology is compliant with prevailing privacy and data protection law and can you motivate why?*
I don't think this technology violates the current privacy law. All the data being used is already public for all to see. Drawing a conclusion from this is not illegal and is currently already often done by people.
- *In which way can you imagine a future impact of the collection of personal data?*
This technology does not store any data. But in general, it is very important that personal data is not leaked and does not end up with people for whom it is not intended.
- *Now that you have thought hard about privacy and data protection, what improvements would you like to make? List them below.*
This technology does not store any data. So I have not come up with any further additions to the technology.

Transparency

- *(How) is explained to the users how a technology works and how the businessmodel works?*

If this technology can actually be used through a website. I can imagine there being an 'about' page explaining how the technology works. In addition, the notebook used to develop the technology is publicly available on github. People with knowledge about machine learning can then see for themselves how this technology was developed.

- *Is it possible to file a complaint or ask questions/get answers about this technology?*

If this technology can actually be used through a website. Then there will also be a 'contact' page. Otherwise, people can look up my contact information on github.

- *Is the technology (company) clear about possible negative consequences or shortcomings of this technology?*

There will be a clear warning that this technology can only be used for predictions. And that the prediction may not be accurate.

- *Now that you have thought hard about the transparency of this technology, what improvements would you like to make? List them below.*

1. A terms of use stating that the technology should not be used for negative objectives would be a good addition.

Inclusivity

- *Will everyone have access to this technology?*
Yes, if you have a computer you basically have access to this technology.
- *Does this technology have a built in bias?*
no, as far as can be estimated in advance, this technique has no biases built in.
- *Does this technology make automatic decisions and how do you account for them?*
Yes, this technology makes automatic decisions on its own. This is handled by giving a clear warning that a prediction should not be taken as truth.
- *Is everyone benefitting from this technology or only a small group? Do you see this as a problem? Why/why not?*
Only users of the technology will benefit. I don't see this as a problem, because I don't expect non-users to be disadvantaged by this.
- *Does the team that creates the technology represents the diversity of our society?*
No, this project is made by one person. This automatically means that an individual does not reflect the diversity of our society. But I don't think this is not a problem.
- *Now that you have thought hard about the inclusivity of this technology, what improvements would you like to make? List them below.*
There are no improvements I would like to apply to this technology in terms of inclusivity.

Sustainability

- *In what way is the direct and indirect energy use of this technology taken into account?*

I expect that the energy consumption of this technology is so low that it is not taken into account at the moment.

- *Do you think alternative materials could have been considered in this technology?*

Yes, this technology is developed on a computer with the programming language python. According to an article on thenewstack by David Cassel, python is the second last energy-efficient programming language from that study. So a different programming language will probably make the total energy consumption lower.

- *Do you think the lifespan of this technology is realistic?*

Yes, I think this technology can be regularly improved or extended by adding multiple new reviews.

- *What is the hidden impact of this technology in the whole chain?*

I don't think there is a hidden impact of this technology.

- *Now that you have thought hard about the sustainability of this technology, what improvements would you like to make? List them below.*

Using a different programming language probably would be better for the environment. But most of this semester is based on using python. Because of that, I will not change the programming language I work with.

Hateful and criminal actors

- *In which way can this technology be used to break the law or avoid the consequences of breaking the law?*

This technology cannot be used directly to break the law. However, this technology can be used to develop sympathy. When someone has done something criminal and he/she can show with a (fake) report, developed with this technology, that the average opinion is positive. So then he/she can create a (fake) sense of sympathy.

- *Can you imagine this technology being used to cross personal- or societal boundaries?*

In some extreme cases, this might be possible. People can quickly use average opinion as an argument. This could cross certain boundaries. But I don't expect this to have dire consequences.

- *Can this technology be used against certain (ethnic) groups or (social) classes?*

Yes, when a classification is done using only the opinion of certain (etic) groups or (social) classes, then this can also be used against them. But to do this, a selection in the input has to be done beforehand. Someone is then deliberately doing this.

- *In which way can bad actors use this technology to pit certain groups against each other? These groups can be, but are not constrained to, ethnic, social, political or religious groups.*

Someone can give extra power to the opinion of a group by making this an average opinion. Using the average opinion may make it seem more likely that a whole group has that opinion. As a result, there may be an even greater difference between the 2 groups when the average opinion is used.

- *How could bad actors use this technology to subvert or attack the truth?*

People might use a fake result to make an opinion seem different than it actually is.

- *Now that you have thought hard about how bad actors can impact this technology, what improvements would you like to make? List them below.*

1. Have some kind of warning that this technology should not be used for bad purposes.
2. create awareness that the results of this technology cannot always be assumed to be true.

Human values

- *How does your technology affect the identity of the user(s)?*
By using this technology you can be seen as a geek. Apart from that, I don't think your identity changes after using this technology.
- *How does the technology influence the user(s) autonomy?*
It perhaps changes a user's autonomy to have their own judgment. Now people themselves often have a judgment whether someone meant something positive or negative. But with the use of this technology, this perspective might disappear.
- *What is the effect of the technology on the health and/or wellbeing of the user(s)?*
There is no effect on the health and/or wellbeing of the user.
- *Now that you have thought hard about the impact of your technology on human values, what improvements would you like to make? List them below.*
There are no improvements I would like to apply to this technology in terms of Human values.

Future

- *What could possibly happen with this technology in the future?*
The technology can become much more precise. Whereby the result can be taken more and more seriously. To a point that people start believing in it. In the same way that people also take a weather forecast for truth, and get angry when it is not true.
- *Sketch a or some future scenario (s) (20-50 years up front) with the help of storytelling. Start with at least one utopian scenario.*
In the ideal world, this technology is used to make life a lot easier. Quickly knowing the general opinion of others can be positively applied in many different ways. Products are improved and produced faster. Politicians can act faster and do what the average citizen wants. And news that incorporates an opinion is made even faster.
- *Sketch a or some future scenario (s) (20-50 years up front) with the help of storytelling. Start with at least one dystopian scenario.*
In a dystopian scenario, this technique is used only for evil purposes. Groups of people are more quickly pigeonholed by the expression of an opinion. People need to think more carefully about how and if they express their opinions. The political system may collapse because all decisions are based on people's average opinions. All opinion pollsters will be out of work and this technology will only do further damage.
- *Would you like to live in one of this scenario's? Why? Why not?*
The first scenario seems pleasant to live in and also the most predictable. In general, there are still more good people than bad, and I don't expect this technology to change that.
- *What happens if your technology (which you have thought of as ethically well-considered) is bought or taken over by another party?*
I don't think another party will change much. They might remove certain warnings or terms of use, but I expect most people are objective enough at the moment and can think for themselves about the consequences of making a classification.

Potential Impact assessment conclusion

Properly informing what the consequences may be is very important. If this technology would actually be available as a tool/website, then it is important to inform users properly.

It is also wise to add a term of use that makes it mandatory to name the tool when it is used in a report. Like "According to the Text sentiment classifier, the average result is ..." This will prevent the average opinion from being seen as truth.

These terms of use will also state that it is mandatory to mention that results are a prediction and cannot be used as a fact or argument. The tool is only made to quickly display general information, not for drawing conclusions.

Datasets

I am going to use 3 datasets that are as different as possible. To keep the training input as variable as possible, I searched for 3 different topics with different structures.

Movies

This is a dataset consisting of 1000 positive and 1000 negative reviews of movies. Unfortunately, it is not known which films the reviews are about. What is known is that they were scraped from imdb.com.

Link: <https://www.kaggle.com/nltkdata/movie-review>

Each row contains a single sentence of a review. This results in very short review texts.

Hotels

This dataset contains 515,000 customer reviews and scoring of 1493 luxury hotels across Europe. The data was scraped from booking.com. Each row consists of very much review and hotel information. This is very interesting for further analysis!

There is one thing unique about this dataset compared to the others. Each reviewer had to fill in a text field for both positive and for negative review text.

Link: <https://www.kaggle.com/jiashenliu/515k-hotel-reviews-data-in-europe>

Because some visitors were very satisfied or unsatisfied, they did not leave a negative or positive review. Or they filled this in as; "nothing positive to report". These kinds of values need to be filtered out. This is something that requires extra attention.

Amazon products

There are very many different Amazon review datasets available. I chose the "Toys and Games" category because it doesn't have much to do with the other topics.

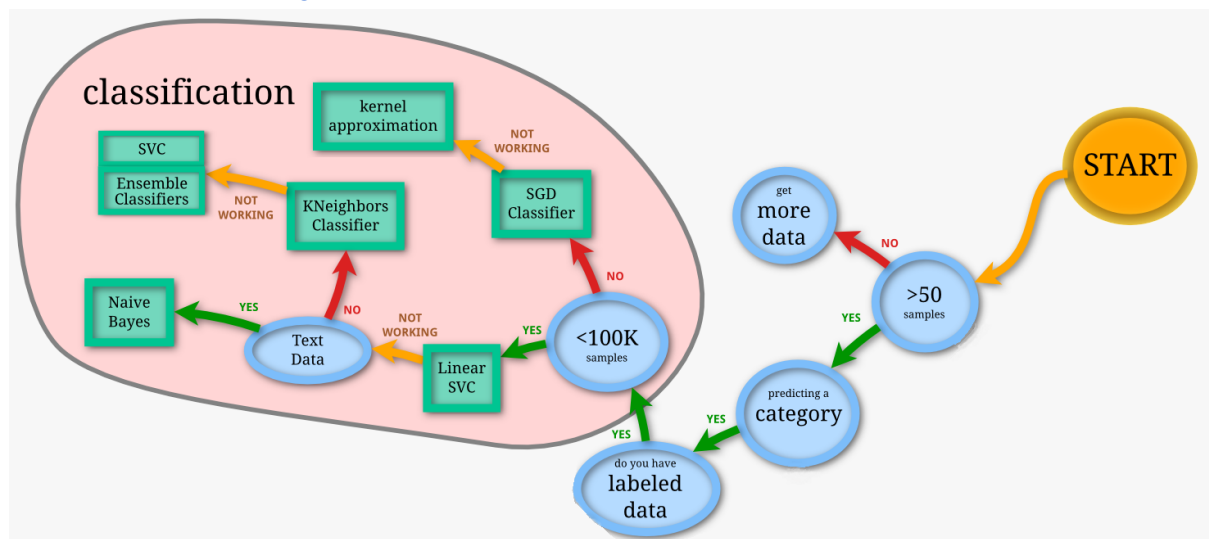
The toys and games dataset consists of 1,828,971 unique reviews, across 78,772 unique products. The data was scraped directly from amazon.com.

Link: <https://nijianmo.github.io/amazon/index.html>

This dataset has not yet been directly labeled positive or negative. But consists of a star rating from 1 to 5 stars. So we need to find a good and fair way to convert this rating to labels.

Modelling

There are very many ways you can classify. A good start is to start by looking at the cheat sheet of [Scikit-learn.org](https://scikit-learn.org)



When I look at my options I come to two things: Linear SVC and Naive bayes.

Linear SVC

Linear svc is a straightforward way of classifying. I am familiar and comfortable with it. I do want to look further into other kernels beyond linear. There are many more options and I would like to research and test with them.

Naive bayes

I have very high hopes for Naive bayes. After all, it is known to work very well with text data. Exactly what this project is about! There are, as with SVC, many different versions of Naive Bayes. To find the best model, I would like to try them all and compare the differences.

Evaluation

At this point, I have learned a lot about the subject. I know what text sentiment means and the difference between positive and negative text. Also, I know what a review is and the different types of reviews.

A clear plan has been made for me and I know what I need to achieve it. However, we have yet to determine when the goal has been achieved.

Domain knowledge verification

How do we know when the model is working well? When the model classifies text sentiment, when will we know if it has done this correctly? This is hard to say because the model is trained with labeled data. This means that the model trains based on the labels in the training data. However, we cannot be sure in advance that these labels are correct. Assuming that these labels are correct, we can make a comparison between the predicted data and test data. I can assume that the model works well when there is little difference between the predicted data and the test data.

Results

We can be satisfied when the model is between 70% and 80% accurate. However, it is important that the difference between the number of false positives and false negatives is as low as possible. This is important because we only have 2 labels.

Deployment

To make the model available to everyone we use anvil. This is a tool that allows you to create beautiful drag and drop UIs and share your local python functions with anyone via a websocket connection.

Literature list

- (2021, 16 March). Review.
Consulted from <https://en.wikipedia.org/wiki/Review>
- (2018, 20 May) Which Programming Languages Use the Least Electricity, by David Cassel
Consulted from <https://thenewstack.io/which-programming-languages-use-the-least-electricity/>