Machine learning

Prediction of wine quality using various machine learning algorithms. The models will train on physio-chemical data derived from samples of wines to determine their quality level.

Created by: MPFGlaser
Created on: May 8, 2022 2:35 PM
Changed on: May 8, 2022 4:56 PM

Context of use: Education
Level of education: Bachelor

Impact on society
What impact is expected from your technology?

What is exactly the problem? Is it really a problem? Are you sure?
The goal of the machine learning model is to predict the quality of a wine (scored 1-10 or good/bad) based on physio-chemical analysis. It can help customers, wholesalers, and wine shops help choose wines to buy, and could help winemakers craft better wines.

Are you sure that this technology is solving the RIGHT problem?
The problem of wine quality can be separated into two distinctive domains: Objective quality and subjective quality. In its current state, the technology can only predict objective quality. However, from my interview, I know that subjective quality plays a huge role in how good a wine is. As such, it would be interesting to look into incorporating this subjective data.

How is this technology going to solve the problem?
The machine learning model will predict the quality of a wine based on several physio-chemical characteristics. Additionally, subjective characteristics could be added to the training data, so that personal taste is also taken into account.

What negative effects do you expect from this technology?
A possible negative effect is that good wines will be classified as bad, and will thus be sold less often than they used to. It could also occur that the model prefers wines grown in an unsustainable manner, and because those wines sell more, more grapes will be grown in that way, impacting the environment through means of artificial fertilisers, pesticides, non-native growing (greenhouses, irrigation, etc.)

In what way is this technology contributing to a world you want to live in?
The technology will enable customers and wine retailers alike to make an assisted choice in buying a wine. The process of finding a good wine would become easier.

Now that you have thought hard about the impact of this technology on society (by filling out the questions above), what improvements would you like to make to the technology? List them below.
Incorporating subjective data, because taste and experience is even more important than physio-chemical data. I would also like to add more data on

the growing process (fertiliser used, pesticides used, etc.)

Hateful and criminal actors
What can bad actors do with your technology?

In which way can the technology be used to break the law or avoid the consequences of breaking the law?
As all data will be anonymous, identity theft is not an issue. However, it could be possible to skew the data by mislabelling the subjective part of the data en masse.

Can fakers, thieves or scammers abuse the technology?
When adding subjective data to the training set, wineries could talk bad about a wine made by a competitor, in hopes of said wine scoring lower than their own.

Can the technology be used against certain (ethnic) groups or (social) classes?
In large groups, subjective tasters could submit lower scored for wines made in certain regions, which they might have a (personal) vendetta against. This could lead to those wines having lower scores, hurting sales, and thus income of the winemakers in that region. A solution to this would be to do blind tasting of the wines.

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.
It could be that wines from a certain region are regarded as better than those from another region. In the extremely unlikely case that people take this personally, it could spark some unrest between those two regions.

How could bad actors use this technology to subvert or attack the truth?
Again, if tasters misreport their findings en masse, it could influence the data and thus the machine learning model's predictions on certain wines, which causes the model to predict lower scores/qualities and could hurt the financials or reputation of winemakers.

Now that you have thought hard about how bad actors can impact this technology, what improvements would you like to make? List them below.
I would add a way of detecting so-called "review bombing" and a method of countering it. Ideally it would not be possible due to diverse testing groups

and data sources.

Privacy

Are you considering the privacy & personal data of the users of your technology?

Does the technology register personal data? If yes, what personal data?

The technology does not register personal data.

Do you think the technology invades the privacy of the stakeholders? If yes, in what way?

The technology does not invade the privacy of the stakeholders.

Is the technology is compliant with prevailing privacy and data protection law? Can you indicate why?

As the technology does not process any personal information, it is in compliance with the GDPR.

Does the technology mitigate privacy and data protection risks/ concerns (privacy by design)? Please indicate how.

The technology mitigates those risks by not using personal data in the first place.

In which way can you imagine a future impact of the collection of personal data?

If subjective data is used, care should be taken to strip it of any identifying details when used alone, or to average it when used with multiple results. As such, there is no way to trace the data back to an individual.

Now that you have thought hard about privacy and data protection, what improvements would you like to make? List them below.

I would make sure that subjective data is anonymised at the source.

Human values
How does the technology affect your human values?

How is the identity of the (intended) users affected by the technology?
The technology could make it so that vinologists and sommeliers are regarded as less important; that is if the model turns out to be very accurate. However, as wine quality is very much decided by personal taste, this would be very unlikely.

How does the technology influence the users' autonomy?
The technology could make certain users feel like they need this to choose a good wine. They could see the score as something more valuable than their own experience.

What is the effect of the technology on the health and/or well-being of users?
When the technology recommends a wine, a user could be more compelled to consume it, moreso than they would normally. Consuming alcohol in large quantities has been proven to be bad for one's health.

Now that you have thought hard about the impact of your technology on human values, what improvements would you like to make to the technology? List them below.
I would make sure that there are clear disclaimers on the judgement of the model and make sure that there are no addicting factors in the application.

Technology Impact Cycle Tool

Machine learning

Stakeholders
Have you considered all stakeholders?

Who are the main users/targetgroups/stakeholders for this technology? Think about the intended context by answering these questions.

Name of the stakeholder
Wine shops/wholesalers

How is this stakeholder affected?
Wine retailers differentiate themselves on selling good quality wines. The higher the quality, the better. If the model recommends them a bad quality wine, it could affect their sales since bad quality wines are not favoured by the customer. In the end, it could influence their financial standing and their reputation both positively and negatively.

Did you consult the stakeholder?
Yes

Are you going to take this stakeholder into account?
Yes

Name of the stakeholder
Wine customers

How is this stakeholder affected?
The public wants to buy a good quality wine. Having a machine learning model/application helping them decide which wine is good could be a helpful crutch to many. In a positive way, it could highlight underrated wines for them, which they might enjoy. On the flip side, it might also recommend a wine to them as good, which in reality is not good at all. This will cause them to question (and rightfully so) the validity and viability of the predictions made by the machine learning model.

Did you consult the stakeholder?
No

Are you going to take this stakeholder into account?
No

Name of the stakeholder
Winemakers

Technology Impact Cycle Tool

Machine learning

How is this stakeholder affected?
Winemakers may want to try and create the best possible wine purely by satisfying the values that the model takes into account. This could cost them a lot of time and effort for very little value, as well as make the wines worse. Alternatively, the outcomes from such a machine learning model could highlight lesser known, underrated wines, boosting their sales and thus income.

Did you consult the stakeholder?
No

Are you going to take this stakeholder into account?
No

Name of the stakeholder
Vineyard owners

How is this stakeholder affected?
As some values are influenced by the type of grapes used, whether or not artificial fertiliser and/or pesticides are used, vineyard owners could change their way of working to better satisfy the model and get a higher prediction. This could result in qualitatively better wines, but also causes a change in the growing process which could financially affect vineyard owners.

Did you consult the stakeholder?
No

Are you going to take this stakeholder into account?
No

Did you consider all stakeholders, even the ones that might not be a user or target group, but still might be of interest?

Name of the stakeholder
Software engineers & UX designers

How is this stakeholder affected?
These stakeholders would have to create an application capable of working with the model, be that user-facing scores, or data entry by laboratories etc.

Did you consult the stakeholder?
No

Are you going to take this stakeholder into account?
No

---

https://www.tict.io   9/16   July 3, 2025 5:01 AM

Technology Impact Cycle Tool

Machine learning

Name of the stakeholder
Laboratories

How is this stakeholder affected?
For the physio-chemical analysis, laboratories have to be used. Use of this model (for training, and further training down the line) requires time, effort, and money.

Did you consult the stakeholder?
No

Are you going to take this stakeholder into account?
No

Now that you have thought hard about all stakeholders, what improvements would you like to make? List them below.
A disclaimer should be shown that the predictions are based purely on physio-chemical makeup, as long as that is the case.

Data
Is data in your technology properly used?

Are you familiar with the fundamental shortcomings and pitfalls of data and do you take this sufficiently into account in the technology?
Wine quality is very much a subjective thing. Objective measurements could be used to get a rough idea of the quality, but ultimately the experience of tasting it will determine whether or not a wine is of good quality.

How does the technology organize continuous improvement when it comes to the use of data?
Assuming the model is (re)trained every so often, it could adapt to the ever changing palate of the population. At certain times sweeter wines will be more popular than acidic wines, and vice versa.

How will the technology keep the insights that it identifies with data sustainable over time?
The current data consists of numerical measurements of physio-chemical compounds present in the wine. Those measurements could be re-taken at any point in time, provided laboratories still exist.
When it comes to the hypothetical subjective data, it is subject to the way of collecting that data. If it is collected locally (i.e. by having an own test/tasting panel) the data has to be stored on our own hardware. If it is scraped from wine rating websites, the data does not necessarily have to be stored by us, as it is already available on the internet. One caveat to this is that the format found on the internet might change over time, requiring work to keep it functional.

In what way do you consider the fact that data is collected from the users?
If the subjective data is gathered from an own test/tasting panel, those tasters would be compensated by means of a salary/compensation. If the data is gathered from the internet, however, users would not have a way to be compensated. As long as the company running this machine learning model is not for-profit, I think this would be fair, as the users have already written their review for the public to see on the internet.

Now that you have thought hard about the impact of data on this technology, what improvements would you like to make? List them below.
I would add subjective data as well as some more objective data.

Inclusivity
Is your technology fair for everyone?

Will everyone have access to the technology?
Anyone could have access to the predictions, but to train the model one would have to be a data scientist.

Does this technology have a built-in bias?
Assuming all data is purely physio-chemical and objective, no. But when subjective data is added, it could be skewed, since people from region A might not like wines from region B as much. Region B would be rated lower in that case, causing a bias.

Does this technology make automatic decisions and how do you account for them?
The machine learning model can only make a prediction on how good a wine is. It cannot and will not fully decide whether or not a wine is bought. It will only assist in the thinking and rating process.

Is everyone benefitting from the technology or only a a small group? Do you see this as a problem? Why/why not?
Everyone - assuming they have some sort of connection with wine - can benefit from this technology. Consumers could get a better wine in their shopping basket, stores could make better decisions on which wine to buy, vineyard owners and winemakers can optimise their processes to make a better wine. All these benefits also have possible detriments; such as a bad suggestion, decreasing sales, higher costs, etc.

Does the team that creates the technology represent the diversity of our society?
The current team consists of one person, so no.
If subjective taste would be taken into account, it would be vital to make sure that the testing group is diverse in all ways possible, though of course you will keep a group of people that are interested in wine.

Now that you have thought hard about the inclusivity of the technology, what improvements would you like to make? List them below.
I would make sure that the testing group is varied in all ways possible.

Technology Impact Cycle Tool

Machine learning

Transparency
Are you transparent about how your technology works?

Is it explained to the users/stakeholders how the technology works and how the business model works?
The data collected can be explained in a white paper, but the exact workings of the model cannot be explained. The technology being machine learning means that it will find links and draw conclusions on its own, without human intervention or a way to explain what exactly those decisions are based on.

If the technology makes an (algorithmic) decision, is it explained to the users/stakeholders how the decision was reached?
It is impossible to explain how the decision was made as the machine learning models do not provide a way to do so.

Is it possible to file a complaint or ask questions/get answers about this technology?
I believe most questions can be answered by reading the white paper that accompanies the end product. However, there will be an FAQ and disclaimers on how the result is only advisory and should not be taken at face value.

Is the technology (company) clear about possible negative consequences or shortcomings of the technology?
Yes. As mentioned above, it should be made clear that the predictions are exactly that. Predictions. They are not in any way meant to replace an experienced taster or your own judgement.

Now that you have thought hard about the transparency of this technology, what improvements would you like to make? List them below.
I would author a white paper explaining the way data is collected and used in order to train the model. (even more extensive than the notebook for the challenge)

Technology Impact Cycle Tool

Machine learning

Sustainability

Is your technology environmentally sustainable?

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

Direct energy is used for training the model(s) and powering the devices that use said models to make predictions. Indirectly, energy is used to gather the information necessary to train the models and do the analysis.

Do you think alternative materials could have been considered in the technology?

No. The predictions and machine learning take place on computers, for which there is no replacement at this time. As for the laboratory analysis, there is a reason as to why the materials they use (such as plastic) are used. They make sure that the analysis is accurate and free of contaminants.

Do you think the lifespan of the technology is realistic?

A machine learning model that is not used is a waste of the energy that was used to train it. However, it will not perish, and could be used in the future.

What is the hidden impact of the technology in the whole chain?

If it so happens that the model systematically points out that wines grown using natural fertiliser and no pesticide score higher, then there is a possibility that the use of those substances will decrease, which would be good for the environment. The opposite is also possible. This all depends on how popular and widely used this technology will get.

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

I would make sure that the end users have a clear understanding of how the model works (on the surface level), and that a thorough, in-depth explanation of the inner workings of the model is not technically possible to give.

Future
Did you consider future impact?

What could possibly happen with this technology in the future?
People could get dependent on the judgement given by the machine learning model and refuse to try something that has not been recommended. As such it would be very hard to get recommended if you are a beginning winemaker or a winemaker who has had struggles with quality in the past.

Sketch a or some future scenario (s) (20-50 years up front) regarding the technology with the help of storytelling. Start with at least one utopian scenario.
In an ideal world the technology could work alongside sommeliers to recommend people the best wines possible. People would use the technology to find the wine that best suits their dinner and have a wonderful experience.

Sketch a or some future scenario (s) (20-50 years up front) regarding the technology with the help of storytelling. Start with at least one dystopian scenario.
People would use the technology as a crutch and get dependent on it, not drinking any wines that were not recommended to them. This would cause smaller, unknown wineries to be forced to close up shop because a wine from an unknown winery in Chile would not be tested or analysed. In the end, only the very big names in wine remain and have an oligopoly.

Would you like to live in one of this scenario's? Why? Why not?
I would like to live in the utopian scenario, as it would be great for someone with little knowledge of wine to have the knowledge of a master at their fingertips.

What happens if the technology (which you have thought of as ethically well-considered) is bought or taken over by another party?
Another company could implement changes that would skew the results in favour of a winery that pays them. As such other wineries would lose business.

Impact Improvement: Now that you have thought hard about the future impact of the technology, what improvements would you like to make? List them below.
I would implement safe guards to make sure that people do not get

dependent on the technology. I would also try to prevent misjudgements by the model causing a big impact in the industry.