1. **Recommendation systems (20 points)**

**MOVIES**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **USERS** | **I** | **II** | **III** | **IV** |
| A | **4** | **3** | **5** | **2** |
| B | **4** | **?** | **5** | **2** |
| C | **3** | **2** | **?** | **4** |
| D | **4** | **1** | **5** | **?** |
| E | **?** | **1** | **2** | **4** |

Green blocks = test sets  
White blocks = training sets  
Orange blocks represent missing rating “?”

My average rating for training set = 3 + 5 + 2 + 4+ 5 + 2 + 3 + 2 + 4 + 1 + 1 + 4 / 12 = 3.

**a.**We know that the overall rating is 3 as shown above, therefore we only need to calculate the biases from the utility matrix:

**USER BIAS =**A = (3+5+2) / 3-3 = 0.33   
B = (4+5+2) / 3-3 = 0.66   
C = (3+2) / 2-3 = -0.50   
D = (4+1) / 2-3 = -0.50  
E = (1+4) / 2-3 = -0.50

**MOVIE BIAS =**   
I = (4+3+4) / 3 - 3 = 0.66  
II = (3+2+1+1) / 4 – 3 = -1.25  
III = (5+5) / 2 – 3 = 2 = 2  
IV = (2+2+4) / 3 – 3 = -0.33

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | I | II | III | IV | Bias |
| A | 4 | 3 | 5 | 2 | 0.33 |
| B | 4 | ? | 5 | 2 | 0.66 |
| C | 3 | 2 | ? | 4 | -0.50 |
| D | 4 | 1 | 5 | ? | -0.50 |
| E | ? | 1 | 2 | 4 | -0.50 |
| Bias | -0.66 | -1.25 | 2 | -0.33 |  |

**b.** Rating = Average + User bias + movie set.

AI = 3 + 0.33 – 0.66 = 2.67  
AII = 3 + 0.33 -1.25 = 2.08  
AIII = 3 + 0.33 +2 = 5.33  
AIV = 3 + 0.33 -0.33 = 3.00

BI = 3 + 0.66 – 0.66 = 3.00  
BII = 3 + 0.66 –1.25= 2.41  
BIII = 3 + 0.66 + 2 = 5.66  
BIV = 3 + 0.66 – 0.33 = 3.33

CI = 3 -0.50 -0.66 = 1.84  
CII = 3 -0.50 -1.25 = 1.25  
CIII = 3 -0.50 + 2 = 4.5  
CIV = 3 -0.50 -0.33 = 2.17

DI = 3 -0.50 -0.66 = 1.84  
DII = 3 –0.50 –1.25 = 1.25  
DIII = 3 -0.50 + 2 = 4.5  
DIV = 3 -0.50 -0.33 = 2.17

EI = 3 -0.50 -0.66 =1.84  
EII = 3 -0.50 -1.25 = 1.25  
EIII = 3 -0.50 +2 = 4.5  
EIV = 3 -0.50 -0.33 = 2.17

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | I | II | III | IV |
| A | 2.67 | 2.08 | 5.33 | 3.00 |
| B | 3.00 | 2.41 | 5.66 | 3.33 |
| C | 1.84 | 1.25 | 4.5 | 2.17 |
| D | 1.84 | 1.25 | 4.5 | 2.17 |
| E | 1.84 | 1.25 | 4.5 | 2.17 |

**c.**

Mse calculation – training set w. raw data:

(3-3)2 + (5-3)2 + (2-3)2 + (4-3)2 + (5-3)2 + (2-3)2 + (3-3)2 + (2-3)2 + (4-3)2 + (1-3)2 + (1-3)2 + (4-3)2 / 12 = 253/12 = 21.083

Mse calculation baseline predicator training set.

(3-2.08)2 + (5-5.33)2 + (2-3.00)2 + (4-3.00)2 + (5-5.66)2 + (2-3.33)2 + (3-1.84)2 + (2-1.25)2 + (4-1.84)2 + (1-1.25)2 + (1-1.25)2 + (4-2.17)2 / 12 = 12.137

(1-12.137/21.083) \* 100% = 42.432% increase.

**d.**

Cosine (**A,B**) = A\*B / ||A||\*||B||   
A\*B = (4,3,5,2) \* (4,0,5,2) = 45

||A|| = ((4,3,5,2)\*(4,3,5,2))  
= sqrt(54) = 7.3

||B|| = ((4,0,5,2)\*(4,0,5,2))   
= sqrt(45) = 6.7

**45/7.3\*6.7 = 0.920**

Cosine (**A,C**) = A\*C / ||A||\*||C||  
A\*C = (4,3,5,2) \* (3,2,0,4) = 26

||A|| = ((4,3,5,2)\*(4,3,5,2))   
= sqrt(54) = 7.3

||B|| = ((3,2,0,4)\*(3,2,0,4))  
= sqrt(29) = 5.4

**26/7,3\*5.4 = 0.659**

Cosine (**A,D**) = A\*C / ||A||\*||D||  
A\*D = (4,3,5,2) \* (4,1,5,0) = 44

||A|| = ((4,3,5,2)\*(4,3,5,2))  
= sqrt(54) = 7.3

||B|| = ((4,1,5,0)\*(4,1,5,0))  
=sqrt(42) = 6.5

**44/7.3\*6.5 = 0.927**

Cosine (**A,E**) = A\*C / ||A||\*||E||  
A\*E = (4,3,5,2) \* (0,1,2,4) = 21

||A|| = ((4,3,5,2)\*(4,3,5,2))  
= sqrt(54) = 7.3

||B|| = ((0,1,2,4)\*(0,1,2,4))   
= sqrt(21) = 4.6

**21/7.3\*4.6 = 0.625**

**e.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **I** | **II** | **III** | **IV** |
| A | **1** | **1** | **1** | **0** |
| B | **1** | **0** | **1** | **0** |
| C | **1** | **0** | **0** | **1** |
| D | **1** | **0** | **1** | **0** |
| E | **0** | **0** | **0** | **1** |

We round up ratings from 3,4,5 to represent those who rate their movies with a high score, and 0 for the rest to represent those who rate with a low score. As there only are two variables we need to worry about it is easier to collect data, as it is more basic.

We calculate the Jaccard by using the formula **|User1⋂User2| / |User1⋃User2|.**

Two value = pairs. |User1 ⋂ User2| / |User1 ⋃ User2 \* 100 = 25%.

A & B = 1/4\*100 = 25%  
A & C = 1/4\*100 = 25%  
A & D = 2/4\*100 = 50%  
A & E = 0/4\*100 = 0%  
B & C = 1/3\*100 = 33%  
B & D = 1/3\*100 = 33%  
B & E = 0/3\*100 = 0%  
C & D = 1/3\*100 = 33%  
C & E = 1/3\*100 = 33%  
D & E = 0/3\*100 = 0%  
  
Some of my pairs got 0%, and the highest score was 50%. No par got 100% similarity.

**f.**

1. **Cascades (20 points)**

**a.**

The Cascade model is a great concept which helps explaining how people are affected by other people’s actions and words. To prove my point on just how easily we are influenced by others, I would stretch as far as saying that human beings are herd animals, as we work great in highly coordinated groups and are designed to pick up on social cues and align our behavior with those around us. These “behavior adjustments” is what we know as “norms”, things like taking off your shoes before entering someone’s house, washing your hands before eating etc. are just a few of many behavioral adjustments which has been lead through the generations.

As we are engineered to think and act this way, we could conclude that there are a general belief that the majority is right, and if a group of people is saying that “x” is correct, you are either more likely to agree or more hesitant to disagree to gain social acceptance. As humans are social creatures we do not like to spend time alone.

To further prove my opinion of humans acting like herd animals, and how easily they are affected we need the help from Solomon Asch. Asch’s experiment consisted of people measuring different lines between two cards. On one of the cards he had three lines, where each line had a different length, on the other card however he had only one line with a fixed length.   
When the participants were asked to show which length of the three matched with the fixed one, only about 3% of the participants would provide an incorrect answer, opposed to the 32% incorrect answer if the participants were surrounded by individuals providing incorrect answers. From this experiment we learn that even though people know or think they know the answer, they get peer pressured by others opinions around them and then depart from their own rational theory to then adopts the theory provided by others. This is one of the many examples of cascades.

Another example of cascades and how people are influenced by others. When a plane successfully land people often clap as a salute the pilots for landing them safely back on the runway. However, this does not happen on every flight. If one person were to start clapping, there would be a lower chance for the others to start clapping as well, as he only is one individual and the majority is not clapping. But if we say 30-40% of the planes population would start clapping, it is more likely for more people to start clapping, and as the amount of people clapping increases the chances for the rest of the passengers to start clapping increases.

**b.**

The musical taste experiment by Salganik and Watt were designed to explain why popularity is both unequal and unpredictable, by manipulating a process at the individual level in a controlled way which led to unanticipated consequences at the collective level.

In order to prove their theory, they to provide proof with relevant research, and to do this they created an artificial market. This artificial market allowed them to observe, modify and predict effects of social influence under controlled conditions. As their goal was to examine and find out how social influence impacted peoples preferences and choices, opposed to how people’s preferences and choices would naturally develop without any impact they needed a controlled environment to experiment on. With the controlled artificial environment set up, they started testing on how social influence could both lead to inequality and unpredictability of success.

Their controlled environment consisted of a musical taste experiment where they had picked out 48 songs from up and coming artists and put it on their own website. The website was limited to participants who took part in this experiment. To start of the participants had to take a short survey where they filled in some personal information, agreed to the terms of service and read a guide. They were informed that on the website they could listen, rate and download songs if they so choose to. However they could only download a song after they had listened to it and rated it. The rating scale consisted of a numeric scale from 1 – very bad to 5 – very good.

The controlled environment consisted of eight separate independent evolving worlds which would be in a social-influenced condition, as well as having one independent world. The eight social influenced groups were made for the researches to test out their conditions and see how the worlds were affected by these interferences. The independent world would consist in a independent condition, were no interference were allowed by the researches as a way to gather data on how the results would naturally develop. By having the independent world naturally develop they could then compare their results from the social influenced worlds with them. Each world had an equal amount of people in them, so the results would be as accurate as possible. These “worlds” were in fact webpages set up in a controlled enjoinment, we only call it that because it is easier to think of them as parallel worlds.

They first carried out two experiments, were both of them would have 48 songs, with 8 influence groups and one independent group. However, the difference between them would be that experiment 1 list songs in random order, and experiment two would list songs by popularity.

The results from experiment 1 and 2 suggested that variations in which they measured success were all different from world to world. A comparison between the two experiments highlighted a surprising effect of social influence. When they increased the social influence by changing from song menu form an unorganized grid to a ranked list, the inequality of outcome increased; resulting in “winners” becoming easier to distinguish from the rest of the artists, meaning that the popular artist become even more popular compared to the rest. At the same time the increased social influence also decreased the predictability, as different artists who became “super stars” were all different throughout their eight parallel worlds.

With these test results it is natural to be under the impression that the increased inequality of success; would suggest an increase in importance of skill. Imagine as the popular artists become even more popular, it is rational suspect that the skill is the factor behind their success. However, when the researches took a look across these 8 parallel worlds they found out that luck had become more important.

We can back up this theory if by proving that if skill were to have been more relevant, the same artists would have been the ones who became more relevant/successful throughout all the worlds. But as each parallel world has its own unique result in which artists were considered the most successful, it is only natural to conclude with this result.

Salganik and Watt constructed a measure of average success for each song, by averaging the market share for each song across the eight parallel worlds. They found out that the if they would have had thousands of parallel worlds so they had a lot of data, they might could have concluded that the best songs become the most successful. But as they only managed to gather a limited amount of participants this would not be possible, and from their data they found out that in any particular realization the best song were not particularly the most successful ones.

They did the experiment over with a different population, as their experiment which consisted mostly of males in their teens and 20s. So they scouted around and invited people from around the globe with different ages, background, ethnicities etc. to see how accurate their previous result had been. Even with the change of participants the results were somewhat the same, so to find out if their theory was successful or not they went on with their last experiment.

In their last experiment of self-fulfilling prophesies, Salganik and Watt found out through their artificial worlds that if participants were convinced that a song were “better” than another, their behavior reinforced that pattern. They compared the results from the individual world with the influenced worlds. In the influenced world they switched ranking between a good song and a bad song. By giving the bad song a high rating, the users behavior recognized the “bad” song as a “good” song, and therefore it somewhat kept its rating.

To better understand what was written above, it can be explained by Merton (1948), who said that a “The self-fulfilling prophecy is, in the beginning, a ***false*** definition of the situation evoking a new behavior which makes the original ***false*** conception come ***true***”. Merton, Robert K. (1948), "The Self Fulfilling Prophecy", Antioch Review, 8 (2 (summer)).

Salganik and Watt’s experiment about information cascades just show how easily we are influenced by others actions and opinions. Their experiment showed that when participants were tricked into believing a song was “good” just by looking at the amount of downloads and likes left by others. With data gathered from this experiment, we can better understand the “collective mind” like behavior and it probably helped a lot of people better understand the point of marketing. Marketing is basically manipulating people into liking or buying a product

**c.**

We learn from Cass Sustein in chapter 4 of Republic that cascades are a product of social media. In this day and age information, as well as false information can be spread to hundreds of millions of people with the press of a button. It is near impossible to predict cascades, but they happen all around us, whether that is when sharing your interests, hobbies, thoughts and opinions through them. We say that cascades are driven by information, whether that is for the good or worse it is up to the individual to decide. The problem in this day and age is that the flow of information moves so fast, it is hard, if not impossible to know what is true and false, and where this information emerged from. A big thanks to this is the evolution of cellphones, computers and the internet. As we do not have access to this direct information, we rely on others that we trust through their statements or actions.

To help us understand the social dynamics behind cascades, Sustein distinguished it between two kinds of cascades: Informational and reputational.

The principle behind an **informational cascade**, is when people with a small amount of information about a particular topic, stop relying on their own limited gathered information and instead adopts the beliefs of others, often people that they trust. But in times where social media is so predominant a lot of people are influenced online and accept certain beliefs simply because they think it is what others believe.   
An example on how easily it is to be affected it if you visit forums like “Reddit” (upvote/downvote/commentbased webpage) and visit their “SubReddits”; which is a forums dedicated to a specific topic on the website. It is easy to get persuaded on these subreddits, as they contain a specific topic and often have up-votes in which people agree or down-vote disagree with them and their opinions. A lot of up-votes on a post or comment might persuade you to think that they are right etc.

A **Reputational cascades** adopt popular beliefs in order to gain social acceptance or to avoid disapproval amongst peers. If everyone have a similar theory, people tend to find it harder to go against that theory. However new knowledge and information can change how people look at a theory. The reputational cascades can best be explained with the tale “The emperor’s New Clothes” by Hans Christian Andersen. In this tale two weavers promised an emperor a new sit of clothes they said was invisible to all of whom were stupid, incompetent or deemed unfit for their position. While in reality they made no clothes, they made everyone believed the clothes were invisible and when the emperor paraded around the country, no one dared to say they did not see any clothes, in fear of being recognized as stupid or inferior. At last a child cries out “but he is not wearing anything” and with this new piece of information making people more confident that the emperor is truly disrobed and not wearing any “invisible clothing”.

The child crying out help show how fragile cascades really are. As any public or precise information can change the actions, as well as the directions of cascades. I do agree with Sunstein that most people have a herd behavior, where if represented with minimal information, the individuals tend to imitate others on the premise that the majority cannot be wrong.

The general problem is when internet users start to pass on information they assume to be true, but cannot possible know to be true based on the information gathered from others. Some people’s inability to think critically of information presented to them is why things like fake news and anti-vaxxers exist. When someone is that easily influenced by groups of people, to the abandon any independent thoughts, is when problems arise.

Although a lot of information passed over social media and the internet may be harmless and constrain truthful information it is hard for many to differentiate the two. When people in cascades take in information online and does not disclose and does not disclose where they received the information when passing it along, confusion and misinterpretation occur. Constant exposure from groups with similar views can also put fuel into this mix.

Sustein then warns us of the danger of technology allowing people to wall themselves off from others, are a risk for the people involved as well as society as a whole. I think what he means by this is that if no one is there to come with counterarguments and make them think of what opinions they have, it is easy to be misled and start to believe something which most people are aware is false from the very start.

The weaknesses in Susteins argument is our ability to sense out false information, fake news and scams. When reading his book you get the feeling that he is talking about the general population and just how easily they are manipulated and influenced. What he fails to address is that most people are very capable of thinking critically. When something seems to good to be true, it most probably is, and therefore people will avoid it.   
One of the reasons why fake news gained such tractions, is that a lot of people got their news from Facebook, and thought that if something was “fake”, Facebook would simply remove it. They left their own responsibility of thinking critically because they trusted Facebook too much, however the ones affected by fake news were a minority, as most Americans saw through the fake articles.

I agree that cascades operate in social media, but I would argue that people argue and come with counter arguments/ideas all the time. There are very few places if any where people agree 100% on things. There will always be contradictions in which people are opposed to each other and would therefore argue. This argumentation would then bring in new information and would invite to new though processes, which could help break the “fragile herd behavior”.

1. **Cambridge Analytica (20 points)**

**a.**

Cambridge Analytica was a data- mining, analysis and brokerage company, which used collected data to predict and change audience behavior. The company collected their data through an app called “thisisyourdigitallife”, built by academic Aleksandr Kogan, to gather basic profile information of users on Facebook as well as what these users “liked” on the platform to create some sort of mapping over behavioral patterns within groups of people.

Although, only about 270,000, where most of whom were paid a small amount to enable this Facebook application, consented to having their data harvested, the app also collected information of the test-takers Facebook friends leading to an accumulation of over 50 million raw profiles.

Kogan managed to collect all this data by abusing one of Facebooks platform policies, which allowed app creators to collect and access information from friends of their app-users as long as the data was spent to improve the user experience. Facebook prohibits any kind of data to be sold or transferred to any ad-network, data broker or other monetization related services.

Kogan then downloaded all the data he collected to his private database, which he then later forwarded it to Cambridge Analytica, violating Facebooks policy and the app agreement which stated that all user data collected in this Facebook app were to be used for academic purposes only.

Cambridge Analytica then used the data forwarded by Kogan, to accurately predict by using a factor model similar to the one Netflix uses to predict movie reviews. By analyzing things like Facebook “likes” they could distinguish between people’s genders, race, sexuality and political view with an up to 95% accuracy just from these like patterns.

“The algorithm and database together made a powerful political tool. It allowed a campaign to identify possible swing voters and craft messages more likely to resonate”- The Guardian. With Cambridge Analytica having significant ties to some of President Trump’s most prominent supporters and advisers, the Trump campaign had no issue with getting their hands on this processed data and used it to target and create online digital advertisements in favor of their political party. Most of the ads posted in favor of the Trump campaign were posted on Facebook and investigation also shows Russian interference in the election, by spreading fake news and propaganda in favor of the conservative party by abusing Facebooks news feed.

The combination of the rise in Fake news and political ads in favor of Trump, were said by experts to have played a massive role in the election result. Further investigation shed light on Facebooks lax rules around the sale of political advertisements, as Russian propaganda spreaders could literally pay with Russian rubles for American political ads with no questions asked, as political ads online are not regulated as closely as they are on TV or radio.

News sources and foreign countries then criticized Facebook’s poor effort in removing Fake news on their platform and thus creating confusion among voters as well as shedding light on how much influence the social media provider has over people and their political views. Facebook is then said to have updates their policies and tightened their rules around advertisements sold on their platform. Anyhow Facebook is said to have lost billions of USD in stock value after this incident and has since then removed over 1.2 billion fake users.

**b.**

Even after Facebook updated its “terms of service” after the new GDPR regulations, they are still operating in the legal gray area within the regulations. In the new TOS Facebook denies any access to Facebook until the user opt-in for ad targeting, cookies and various other data processing purposes.

Attempting to collect consent in this manner, I would argue to be unlawful as it breaches with article 7 (1); freely given, non-conditional consent. When I read more about the GDPR, I also think that they also could break principle 1; the act lawfulness, fairness and especially transparency; as their TOS is hidden behind walls of text, and you could call simplified or specified, which the GDPR specifically asks for. And if users do not comply with the TOS consequently they get denied service.

Facebook has defended cookies and said “Cookies help us provide, protect and improve the Facebook Products, such as by personalizing content, tailoring and measuring ads, and providing a safer experience”( <https://www.facebook.com/policy/cookies/>).

I would argue that the explanation around given around cookies is too vague to be counted as valid within the GDPR jurisdiction. They avoid explaining reasoning behind cookies cookie use by explaining the term it with broad and non-specific consent that cover a variety of the cookies activities.  
Saying that their cookies help “creating personalized content”, “providing a safer experience”. Is just too vague in my opinion and it does not help us find out what specific purpose the data collected is for. GDPR also want instances where personal data processed, to be explicit and legitimate, with determined proof of time collected and for the data to be adequate, relevant and limited to what is necessary for the purposes for which they are processed.” (<http://www.privacy-regulation.eu/en/recital-39-GDPR.htm>).

Shortly after GDPR went into action Facebook had a security breach, which lead to roughly 50 million accounts having their information leaked. Investigation found out the security breach was caused by a vulnerabilities within Facebook’s code which permitted attackers to access and steal private information. In my eyes Facebook failed to operate within the   
6. Principle of GDPR; Integrity and confidentiality as they failed to provide appropriate security within storage of personal data and protecting the data obtained against unauthorized or unlawful processing of their data.

It is hard to point and be specific with what breaks with the GDPR, in such a manner that a large corporation such as Facebook could be fined for, as GDPR is very vague about measures organizations should take. Facebook being the large corporation it is, serving hundreds of millions of people every day it is eventually bound to have mishaps happen such as leaving exploits open in their system or to eventually get hacked. I personally feel like they should consider to some degree to encrypt all personal data gathered, so if a potential breach happens again the information would remain somewhat “safe”, and considering the sensitive data Facebook sits on such as peoples sexual preference, social orientation, interests etc. it would benefit all parties involved.

**c.**

***1. How Cambridge exploited the Facebook Data of Millions | NYT***

1,761 likes | 55 dislikes = total votes 1816.  
Ranking: 1761/55 = 32.018

***2. Facebook’s Cambridge Analytica scandal, explained | The Verge***

11,171 likes | 445 dislikes = total votes 11616.  
Ranking: 11171/445 = 25.103

***3. Cambridge Analytica Uncovered: Secret filming reveals election tricks | Channel 4 news***

25,594 likes | 852 dislikes = total votes 26,446.  
Ranking: 30.039

***4. Cambridge Analytica – The Power of Big Data and Psychographics | Concordia***

3,882 likes | 391 dislikes = total votes 4,273  
Ranking: 9.928

***5. What is the Cambridge Analytica Scandal? | The Guardian***

2,351 likes | 321 dislikes = total votes 2,672  
Ranking: 7.323

**Total** video **likes** = (1761 + 1171 + 25594 + 3882 + 2351) = **34 759.**

**Total** video **dislikes** = (55 + 445 + 852 + 391 + 321) = **2064.**

**Average** video **likes** = (34 759/5) = 6958.8 = **6959.**

**Average** video **dislikes** = (2064/5) = 412.8 = **413.**

**Average rating** = 32.018 + 25.103 + 30.039 + 9.928 + 7.323 / 5 = **20.8822**

**Video *1. How Cambridge exploited the Facebook Data of Millions | NYT*1:** 34759 \* 20.882 + 1816 \* 32.018 / 34759 + 1816 = 21.434

**Video *2. Facebook’s Cambridge Analytica scandal, explained | The Verge*2:** 34759 \* 20.882 + 11 616 \* 25.103 / 34759 + 11616 = 21.939

**Video *3. Cambridge Analytica Uncovered: Secret filming reveals election tricks | Channel 4 news*3:** 34759 \* 20.882 + 26 446 \* 30.039 / 34759 + 26446 = 24.838

**Video *4. Cambridge Analytica – The Power of Big Data and Psychographics | Concordia*4:** 34759 \* 20.882 \* 4273 x 9.928 / 34759 + 4273 = 19.682

**Video *5. What is the Cambridge Analytica Scandal? | The Guardian*5:** 34759 \* 20.882 + 2672 \* 7.323 / 34759 + 2672 = 19.914

From my calculation we can conclude that Channel 4 had the better score at 24.838. It is hard to say if a video is “better” than another, as every person has different preferences. Although most of the videos are written and made by professionals, it is hard to choose a video which is better. Even though the ranking were close, I would argue that Channel 4 had the best video. Not only did it have the best score, but also the most votes registered at 26 446.   
The more votes a video have, the more accurate the results are.

I can prove this concept we could explain it through an example. If someone were to have one up-vote on a video and no down-votes, it does not represent how well made the video was, as there is only the opinion from one guy and it does not accurately represent the score of the video.

**SOURCES:**

Cass R. Sunstein. (2017). #Republic: Divided Democracy in the Age of Social Media. Princeton University Press. Chapter 4: Page 95-129.

Christopher G. Brinton & Mung Chiang (2017): The Power Of Networks. Princeton University Press. Chapter 9, Chapter 7, page 123.

Matthew J. Salganik & Duncan J. Watts: ‘Social Influence: The Puzzling Nature of Success in Cultural Markets. (Article provided by UIB).

Merton, Robert K. (1948), "The Self Fulfilling Prophecy", Antioch Review, 8 (2 (Summer))

**LINKS:**

<http://www.niemanlab.org/2018/03/this-is-how-cambridge-analyticas-facebook-targeting-model-really-worked-according-to-the-person-who-built-it/> [06.11.18 – 10:15 opened]

<https://www.theatlantic.com/technology/archive/2018/03/the-cambridge-analytica-scandal-in-three-paragraphs/556046/> [06.11.18 – 11:01 opened]

<https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html> [06.11.18 – 13:24 opened]

<https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> [06.11.18 – 16:40 opened]

<https://no.wikipedia.org/wiki/Cambridge_Analytica> [06.11.18 - 21:30 opened]

<https://eugdpr.org/the-regulation/> [06.11.18 - 21:29 opened]

<https://www.facebook.com/policy/cookies/> [10.11.18 – 13:42 opened]

<http://www.privacy-regulation.eu/en/recital-39-GDPR.htm> [10.11.18]

(RESULT I GOT FROM YOUTUBE CAMBRIDE ANALYTICA) [12.11.18 – 11:27 am opened]