Oppositionsprotokoll

Rapportförfattare
Richard Nysäter och Tobias Reinhammar

Rapportens titel
Song Similarity Classification
Using Music Information Retrieval on the Million Song Dataset

Opponent
Rikard Blixt

Var det lätt att förstå vad exjobbet gick ut på? Kommentarer.
To understand that the goal of the report/study was very easy, with a clearly define problem statement as well as a hypothesis. Even from only reading the abstract you had a clear sense

Hur tycker du att titeln avspeglar rapportens innehåll?

Absolutely, the title is a direct match to the contents of the report, the entire report aimed at Classifying Song Similarities using MIR from MSD.

Hur beskrev författaren bakgrunden till exjobbet? Finns det en introduktion till och översikt av området?
While there's a background for the topic, it could have been done better. A lot of times they use a none-standard term (e.g. Timbre) before they explain it. Also major parts of early content in chapter 2 (Method) feels like it fits more into the background/introduction.

Hur väl har författaren motiverat sitt val av metod att angripa problemet?
Not very well, for instance, they never state why they picked the algorithm they decided to use (k-NN). It feels like going from "This is a problem" to "This is the solution" without explaining WHY it is the solution, or even why this algorithm over another one.

Är metoden väl beskriven?

The method is fairly well described, it does however get confusing with lots of new terminology coming up and being explained, it is sometimes easy to miss a very important part, which was hidden in between terminology explanations. There was a few things that were badly explained, mainly why '7' was picked in k-NN.

Har författaren presenterat sina resultat tydligt?

The results are not that clearly presented, with confusing graphs that are hard to understand.

Finner du författarens slutsatser trovärdiga?
I find the conclusion very short and not really saying a lot, the little actual concluding they do, they do it very well and feels credible with rest of the report.

Vad tycker du om litteraturlistan? Vilken typ av litteratur finns med? Förefaller litteraturen relevant?
The references could have been written a lot better, it was hard to see where one part (e.g. authors) ended and the next part started (e.g. title). As far as I can tell no standard (Vancouver etc.) was used. There's also a few mistakes here and there, such as parts of a link is missing or fetch date missing. Additionally the references used in the text were inconsistent, sometimes they were used before a 'dot' and sometimes after. They only used websites, but some are from online books and others are from reports.

Vilka avsnitt i rapporten var svåra att förstå?
Results and the references, for reasons stated earlier.

Övriga kommentarer om rapporten och dess struktur.
Abstract and Sammanfattning doesn't contain the same content. E.g. the abstract talk about "real-world applications" (Swedish: "appliceringar i riktiga världen") while the sammanfattning states "i exempelvis musikapplikationer".

Vilka är arbetets/rapportens starka sidor?
A well explained method and a good interesting discussion are the two best parts of the report.

Vilka är arbetets/rapportens svaga sidor?
Results and the references are the two weakest links in this report, for reasons stated earlier.

Hur bedömer du arbetets nyhetsvärde?
The topic is a fairly hot topic today and will likely continue to be for the foreseeable future, with the online music industry steadily growing thanks to services such as Spotify. In this industry having good "recommended songs for you" can be very important, as it can keep the customers happy, using the service more.

Sammanfatta arbetet på ett par rader.
A well done study and report on a major present day topic. The results were not as good as previous studies, but still managed to show it is likely possible. In about 25% of the songs, the computer and a human thought the most similar song (out of two possible picks) was the same in at least 90% of cases, and in 50% of the songs it was at least 80%.

Frågor:

1. In your discussion you state that your study did not achieve the results of the Aucouturier and Pachet study, and a possible reason is that your study took into account many more variables than the A&P one. Why do you think this is? Shouldn't it be more accurate when considering more variables?
2. If you would have had a larger user base and more songs to do the "Phase 1" test, letting users compare songs, so you have more data for the algorithm. Do you think that the results would have been about the same or significantly better/worse?
3. How did you actually conclude the k-value in k-NN? What you wrote in the study was extremely hard to understand.