I am considering that after I finish this course, I can continue a project that I have done for my Bachelor Thesis for which I processed the audio signal of a collection of songs and obtained a collection of parameters that could help me to group the songs according to their genre.

I had a collection with 300 songs from 6 genres: rock, classical, jazz, disco, country and hip-hop. For a classification algorithm like knn, the parameters had a quite good accuracy, around 70%. I used the resulting parameters for making a "smart music player" that is automatically picking

the next song to b played by using k-nearest neighbour. I was even then thinking to have a way to visualise the placement of the songs in space according to the similarity of their parameters and used a tool RapidMiner that had implemented the Kohonen Maps algorithm that is mapping n-dimensional items to 2d

or whatever dimensionality by approximately preserving the Euclidean distance between the items. I obtained a map much like this: <img>

Now I would like to improve my musical player and somehow display an Interactive Self Organising Map either as desktop application or maybe even on a webpage associated with the application. As it can be seen in the image above, we can see the genre of song that is at a specific location but I would like to

make an interactive map that will also show me details about the song like the name, the author when I click or hover above it. I would also like to maybe see the songs that are chosen by knn to play next. Now that I have seen something similar at the lecture, I am also thinking to maybe, when I click the region

with songs of a specific genre, to also see a split into subgenres of those songs.

Such an improvement to my project can have two types of targets: first of all, it will help me to see the correlation between parameters obtained from audio signal processing of songs based on frequency or other characteristics of audio signal and the genres classification. Since I obtained quite good performance with little refined algorithm,

some improvement will certainly increase the accuracy of making an automatic classification into genres and also accuracy of music player of choosing the next song with similar genre. Secondly, such visualisation will also be great for possible users of the application since they will have an awesome view of their collection of songs and

it will maybe have an use similar to famous Shazam application but this time for getting an interesting placing in genre for new songs that are developed.

There are many sources online to get songs in mp3 format and I am thinking that maybe I can also make into a larger application to have a website and users use their local machines to process the songs on their machines by using the algorithm as I implemented and then to upload the results on a common server and then their new entries would appear

in a big SOM that can be viewed by all users of the application or website. By doing this, people can maybe get a guess of the genre of a new song if they doubt about its classification.

Assignment 2 ex 3.1

As we can see in our visualizations, it is hard to understand how the weather has evolved during the last 100 years and I think this is not much because of the visualizations themselves, but mostly because of the data we have used.

To obtain an improvement I would, first of all, make some statistical analysis on how the weather changed in time. For example, we could keep calculating the mean of temperature during the whole year but also calculate for example variance to see how much has the weather changed during those years. By having this information, I think we could see if the weather became more unstable in the recent year, something expected due to the global warming. I think we could still use the plot as in our first visualization since it is certainly one of the best methods to see how a variable evolves in time and by plotting different statistical analysis maybe we can somehow find what big changes happened to the weather for the past 100 years. We could also calculate the standard deviation for each month according to the mean of the corresponding year and still plot that in our heat map from the second visualization and maybe we can still see valuable information by using the same code we already have.

Secondly, I would also consider data from other weather stations and then provide a way to visualize how the weather has changed in different regions of a country, or a continent. I think for this we could calculate how much has the mean temperature from each weather station changed during the last 100 years and obtain a single value, positive if the average temperature has increased and negative otherwise. By taking these values, I think it could be very useful to make a heat-map like this <http://icmconsulting.com/media/uploads/Geothermal_heat_map_US.png> and colour the regions surrounding each weather station with a colour that corresponds to the values we obtained. By having such a visualization, we can see if the weather became warmer or colder in some regions, or if in some regions the average temperature has increased faster than in other regions. I think this would also provide useful information on how weather in regions with different geographical aspects evolve in time. For example, Sibiu is a town located in an area surrounded by mountains and many forests that have a big impact on keeping the temperature quite constant and provide protection from important winds. Maybe in regions from south of Romania where the land is more plain, with few forests and heavy agricultural activity, the average temperature has increased faster and that would be visible in the type of heat-map that I am suggesting. An issue for making such visualization is that not all weather stations provide consistent data since 1880 as the station from Sibiu but I think there are certainly ways to deal with this problem.

Finally, we all know that weather is very complex and there are several aspects to be considered when analysing how it evolves. Besides temperature there are many other parameters to be considered: wind, humidity, air pressure, number of storms, etc. Even with all these measurements, weather is still hard to forecast. For a complete analysis of weather evolution during the last 100 years, data describing all these aspects should be considered and make similar visualisations for these parameters in a manner to be able to compare the results. For this purpose, I would still use the heat-map and the plot as described previously and I would add some buttons with settings to change what we want to display and have the possibility to view 2 interactions at a time so we can make a comparison. Those buttons should be able to set the parameter we want to view and if we want to get the average for that parameter during one year, only one month or only one season.

I can conclude that we can answer to the question of the assignment by getting data from several weather stations and displaying some statistical analysis for this data in a manner so we can view how the weather parameters over a large map in a manner that we can make comparisons between different settings of the interactive visualisation.