


CHARLES UNIVERSITY
Faculty of mathematics
and physics

Eliška Janásková, M.S. Thesis
Statistical machine learning with applications in music

Superviser Report

June 3, 2019

I have defined the main goals of this work as stated by the project task as following: "This work aims to review the current state of the art in statistical machine learning and apply it to music composition. The thesis should first briefly describe the general methods of statistical machine learning. Second, thesis should survey the existing methods applied in machine generated music."

It is obvious from the project description that the nature of the project was to a huge degree experimental with little expectations for contribution to the theory of statistical machine learning. My own position is that works of the high risk – high reward should be given an opportunity and this is one of such cases. My own expectations about the results have been mixed – we are talking about a project that has already attracted a significant attention from mathematicians, computer scientists and artists, but with a very limited success. To put the state of the art in some broader perspective, the machine learning technology is relatively successful when working with visual data (computers are now better than humans in recognition of faces, they can drive autonomous cars or they can produce artificial real looking photographs), but music composition is still fully a domain of humans at this stage. However, it is conceivable that the computers will soon be able to compose music better than humans.

Thus in order to make this project doable, I have had to define the goals that are within the ability of a student to process within a limited time and produce a written report. My qualified judgment of what is doable is a review of the current machine learning techniques and their applications in producing outputs in terms of computer generated melodies. As writing own music composition code would be far beyond the amount of work expected for a diploma thesis, I have suggested to use the existing platform Magenta from Google that has been recently available to developers.

Both goals were successfully addressed in the thesis. The review part is well done, one should keep in mind that this is simply a survey of the existing methods with no expectation for new results. The overview is written in her own words (judging from the similarities report) in a solid English. As there is a huge ongoing demand for these new methods, I find highly beneficial that the author has made this kind of work. This clearly goes beyond the current coursework.

The main original contribution of this work is in the production the actual compositions. The implementation was done in Python, which itself is a measurable learning benefit. For the music generation task, one needs to produce a well defined training set. Python is able to work with MIDI files, so the training set should be in this format. Producing the training set is not an easy task – my own suggestion was to use a relatively large (in a scale of hundreds of compositions), but homogeneous set. The natural choice are Beatles compositions due to their prolific and well documented (existence of available sheet music with reasonably correct melodies) production. The set produced can be viewed as a canonical training set for any subsequent machine learning projects (limited to Beatles compositions) as it includes their $\sim$200 most known melodies.

Mrs. Janásková additionally trained the computer to produce sample melodies using Google Magenta. This is by no means an easy task – I myself have not been able to make it work on my own systems as it requires some non-trivial tweaks of their code. The thesis discusses the issue of "optimal length" of training which is itself interesting. The results are the MIDI files that accompany the thesis. One can clearly see that the computer generated melodies are no match for the real Beatles, but that was fully expected as any system that could successfully generate artificial Beatles melodies would become an instant global hit, which is not the case. What statistical features make the computer generated melodies different is documented in the thesis. My short impression is that human melodies tend to evolve mostly on neighbouring notes in a given scale with relatively long monotone sequences (making melodies a non-markovian process). In contrast, computer generated melodies jump more and keep less monotonicity which makes it less aesthetically pleasing. This is documented in the thesis by the transition matrix of notes in both the original Beatles and the computer generated melodies. In addition, the length of notes have distinct pattern for both original and computer generated melodies.

In conclusion, I believe that Mrs. Janásková successfully fulfilled the diploma thesis task and that all possible limitations of the work are due to the fact of the difficult nature of the project and the fact that the current state of the art that does not allow to go far beyond what is described in the thesis. There are several possible extensions that could have been included in the thesis – one example is to train the computer to produce different (and short – a small number of notes) alternative endings to some widely known melodies and check the statistical frequency of the resulting endings. A more ambitious addition to the project would be to include a rating algorithm that would evaluate the quality of the composition and thus pro-

duce a filter that would immediately discard low quality generated melodies, but this turned out to be a task far beyond this thesis.

Summary: The thesis satisfies conditions of a master thesis and I recommend that it is **accepted as such.**

Jan Vecer,
KPMS, MFF UK,
Sokolovska 83
18675 Praha 8
Czech Republic
Email: vecer@karlin.mff.cuni.cz