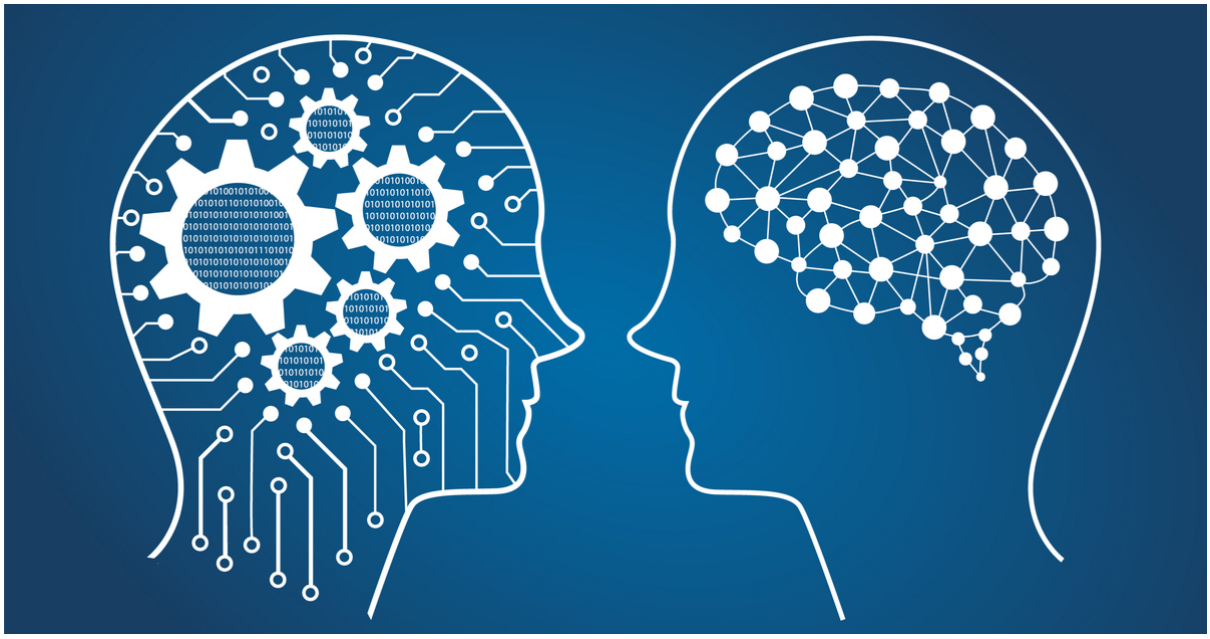
Martin Kaptein 1527038
Klassieke muziek, Zwolle, ArtEz hogeschool voor de Kunsten
Mentor: Rianne Heezen
31 May 2019

Can J.S.Bach be imitated using machine learning?

Research report



1

[1] https://news.sophos.com/en-us/2017/07/24/5-questions-to-ask-about-machine-learning/ (visited on 01.03.2019)

Preface

My research aims to bring two completely different worlds together: The world of technology and the world of music. This is an attempt at understanding the possibilities of machine learning and, by extension, artificial intelligence, in the context of music making.

Personally, I am fascinated by both topics; whilst pursuing a musical career at ArtEz, information technology, more precisely, artificial intelligence, has always been an interest of mine. Now, with my research, I have finally the opportunity to bring these worlds together, while extracting helpful and valuable information.

I am thankful for Ivo van Emmerik, Harrie Janssen, Koen van der Meer, Rianne Heezen and Oscar Ramspek.

Summary

In my research, I have dived deeper into the concept of machine learning, and explored, what capacity it has to learn and recreate works of music: How do experts perceive three different chorales, that have been generated by a machine learning algorithm, with three different adjustments, that has been trained on 389 real chorales by J.S.Bach, in terms of its level of imitation of J.S.Bach? For this research I have used a machine learning algorithm called "DeepBach" and documented in the following paper: "DeepBach: a Steerable Model for Bach Chorales Generation" (Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016).

To complete this research as efficiently as possible I have divided it into methodically three parts: Literature research (to analyse the characteristics of a chorale by J.S.Bach as well as to analyse the working of the above mentioned machine learning algorithm in order to have it generate three different Bach chorales of different complexities), experiment (the step in which three different versions were generated) and, finally, the most important part, the evaluation in the form of interviews with Bach-experts. These experts were to judge the three different results of the machine learning algorithm, and analyse it thoroughly in all possible dimensions. In general, all experts discovered very similar results, and came to, generally speaking, the same conclusion: The three different "fake" chorales, could be, without any doubt, seen as clearly not originating from J.S.Bach. However, the experts also agreed on one specific chorale to be the best in the regard of its imitation to J.S.Bach. Reasons on why this might be the case, will be outlined later in this research report.

These results are very important because they show that human creativity is still a very difficult concept to imitate. However, what it also shows, that, if done right, machine learning has indeed a capacity to make surprisingly "good" advancements.

Keywords

artificial intelligence, machine learning, J.S.Bach chorales,

Table of contents

Preface 2
Summary 3
Keywords 3
Table of contents 4
1. Introduction 5
2. Research method 10
3. Results 12
Results of the Interviews 15
Conclusion 19
Discussion 20
Appendix 23
Reference list 29

1. Introduction

In our modern society, we are surrounded by technology. Everywhere, we look around us, there is technology. But what if technology becomes smart? Will human musicians ever be replaced by technology?

Since a very long time, the idea of a possibility to replicate the creative process of a human brain, has fascinated authors, philosophers and technicians alike. In ancient Greece, there was a god for automation, Talos.[2] In 1863 the author Samuel Butler rose the question of the evolution of artificial consciousness.[3]
The last 50 years, have provided us with a new possibility to approach this problem: It's the advent of the concept of machine learning.[4]

> Machine Learning is the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions.[5]

In my own role as a practising musician, I have often wondered whether the job of composing and performing music will ever succumbed to machines. It is natural, yet incredibly naive to assume, computers will never acquire the ability to mimic the functioning of the human brain: Machine learning, combined with the use of neural networks does exactly that.[6]

As Information Technology has also always been an interest of mine, since my early ages, I have decided to use my knowledge, in this research. Hence, I will as a matter of fact combine the topics of classical music and information technology.

---

[2] https://en.wikipedia.org/wiki/Talos (visited on 25.02.2019)

[3] "Darwin among the Machines". The Press, Christchurch, New Zealand. 13 June 1863.

[4] R. Kohavi and F. Provost, "Glossary of terms," Machine Learning, vol. 30, no. 2–3, pp. 271–274, 1998.

[5] https://emerj.com/ai-glossary-terms/what-is-machine-learning/(visited on 24.02.2019)

[6] https://en.wikipedia.org/wiki/Artificial_neural_network (visited on 25.02.2019)

My research will be taking place in the context of my study in the ArtEz conservatory, as an active, thinking, musician. I have used my contacts at the ArtEz conservatory. However, I will also present my knowledge in information technology, as I have not only access to literature and software on this topic, but I am an active programmer.

Attempting to shed light on the topic of abilities of computers to mimic the creative process of a human is not only philosophical, but existential. As a matter of fact, any musician should be interested in the changes of the environment she or he acts in.

These are the reasons I have chosen this topic.

Theoretical context

The topic of machine learning is a very widely discussed topic: One look at the wikipedia page of machine learning[7], already shows a wide variety of types of machine learning algorithms and models. However, what I am, more precisely, interested in, is the way to generate a Bach chorale using machine learning, for a result that best imitates Bach. For this problem also exists literature, which I will elaborate on in the following.

What is also important to note, is the fact that music generation algorithms exist already for a long time: The difference is the fact that many of those algorithms are rule-based. A popular example is WolframTones[8], which works with pre-programmed mathematical models. Rule-based means, that the rules of composition are programmed by hand, by a human.

An example of the latter, which deals with precisely Bach chorale generation and imitation is a solution proposed by Kemal Ebcioglu in 1988[9]: His solution was to create a rule-based expert system with more than 300 rules to generate a chorale in the style of Bach. This is known to have generated results which completely imitate the sound of Bach, i.e. which are completely flawless. Rule-based expert system means that rules concerning harmony, rhythm

[7] https://en.wikipedia.org/wiki/Machine_learning (visited on 26.02.2019)

[8] http://tones.wolfram.com/about/how-it-works/(visited on 23.02.2019)

[9] Ebcioglu, Kemal. An expert system for harmonizing fourpart chorales. Computer Music Journal, 12(3):43–51, 1988. ISSN 01489267, 15315169. URL http:// www.jstor.org/stable/3680335 (visited on 03.01.2019).

and all other musical parameters, have been programmed manually, which took a considerable effort.

This approach, however, doesn't suit the needs for my research, because it doesn't make use of any machine learning. A computer just follows strict rules, as opposed to acting in a more human-like way (which I referred to in my introduction) of learning abstract context:

The approach of machine learning towards music generation, is completely different, as it not rigidly programmed software (which Kemal Ebcioglu refers to as "rule-based expert system"), but rather software, that learns based on a huge amount of input data (usually it is referred to as 'training data').[10] A machine learning algorithm doesn't need to be explicitly programmed to do a certain thing, it finds that out itself. Hence, it is important to distinguish between a human programming a computer, and a computer learning itself.

A more recent approach, which approximates my needs much better is the BachBot model.[11] This approach uses machine learning to generate chorales in the style of Bach. However, this algorithm has also its limitations, mainly because it doesn't allow for a flexible approach. It is important to be able to tweak the machine learning algorithm, in order to get different results, that will later be judged by the committee of experts.

A paper, which I will extensively use is "DeepBach: a Steerable Model for Bach Chorales Generation"(Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016)". This paper provides exactly the approach to Bach chorale generation, that I will need for my research.

Since, for my research, I am more interested in the approach of machine learning, rather than the approach of rule-based Bach chorale generation, I will need to make use of a framework, which uses machine learning, yet which also allows for the modification of certain parameters (which the paper, I mentioned above, refers to as 'steerable'). The above mentioned paper, depicts exactly this kind of algorithm.

---

[10] James, Gareth (2013). An Introduction to Statistical Learning: with Applications in R. Springer. p. 176. ISBN 978-1461471370.

[11] Liang, Feynman. Bachbot. https://github.com/feynmanliang/bachbot (visited on 29.11.18) , 2016.

This is important, because, I will need to generate different examples of chorales, in order to analyse them, with the help of experts, in regards to their level of similarity to J.S.Bach.

I couldn't find any research about the topic of how convincing the results of experiments which attempted to imitate chorales by J.S Bach using machine learning were. This is, what my research will be about. This is the reason, why it is interesting to approach such a technical subject from the point of view of a musician.

Last but not least, I will need some other literature: Mainly I will need literature, which depicts original Bach chorales.

For the latter, I will use:

"Bach, J.S. 389 Chorales (Choral-Gesange): SATB (German Language Edition). Kalmus Classic Edition. Alfred Publishing Company, 1985. ISBN 9780769244204. URL https://books.google.fr/books?id= U1-cAAAACAAJ".

Since there is debate about what exactly constitutes a chorale by J.S.Bach, and there even is the debate about which chorales are actually compositions by J.S.Bach[12] I have decided to constrain myself to the above mentioned source of 389 Bach chorales.

To describe possible differences in terms of counterpoint/harmony:

Laitz, Steven G. (2008). The Complete Musician (2 ed.). New York, NY: Oxford University Press, Inc. p. 96. ISBN 978-0-19-530108-3.

Also, it will be important to have access to literature, which describes various machine learning techniques. For this, I will use:

[12] Carl Philipp Emanuel Bach and Johann Friedrich Agricola. "Bach's Nekrolog" (full title: "VI. Denkmal dreyer verstorbenen Mitglieder der Societät der musikalischen Wissenschafften; C. Der dritte und letzte ist der im Orgelspielen Weltberühmte HochEdle Herr Johann Sebastian Bach, Königlich-Pohlnischer und Churfürstlich Sächsicher Hofcompositeur, und Musikdirector in Leipzig"), pp. 158–176 in Lorenz Christoph Mizler's Musikalische Bibliothek [de], Volume IV Part 1. Leipzig, Mizlerischer Bücherverlag, 1754 – pp. 167–168

Bishop, C. M. (2006), Pattern Recognition and Machine Learning, Springer, ISBN 978-0-387-31073-2

Definitions

Algorithm:
In mathematics and computer science, an algorithm is an unambiguous specification of how to solve a class of problems. Algorithms can perform calculation, data processing and automated reasoning tasks.[13]

Machine learning:
Machine learning (ML) is the study of algorithms and mathematical models that computer systems use to progressively improve their performance on a specific task.[14]

Bach chorale:
A chorale by Bach as found in my list of literature, that depicts the chorales by J.S.Bach I will use specifically for this research.

Research problem and Goal

The goal of my research is to find out the possibilities of computers, more precisely, a machine learning algorithm in the context of approximating real chorales by J.S.Bach as

[13] https://en.wikipedia.org/wiki/Algorithm (visited on 3.12.18)

[14] https://en.wikipedia.org/wiki/Machine_learning (visited on 3.12.18)

judged by experts. In other words, the goal of my research is to find out whether J.S.Bach can be imitated using machine learning. It is important to be aware of the capabilities of modern technologies, even in the field of classical music. My research will provide insight into the latter.

Main research question

How do experts perceive three different chorales, that have been generated by a machine learning algorithm, with three different adjustments, that has been trained on 389 real chorales by J.S.Bach, in terms of its level of imitation of J.S.Bach?
To answer this question, it is necessary to structure it in the following sub-questions:

Sub-questions

1. How is a Bach chorale characterised?
2. How does the machine learning algorithm from the paper "DeepBach: a Steerable Model for Bach Chorales Generation"(Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016) work?
3. How can the machine learning algorithm be adjusted for (three different) results, that approximate chorales by J.S.Bach as close as possible?
4. How do experts perceive differences between a real Bach chorale and three different results of a machine learning algorithm?
5. Which algorithm adjustment yields the best replication of a chorale by J.S.Bach in the perception of experts and why?

2. Research method

The research that I will do will be the qualitative type. This means, that I am going to interview a small group of people. This is because, for this research, I am mainly interested in very detailed answers. Thus, a committee of three people, who have a slightly different

musical background and whom I will interview is more than enough, to judge the outcomes of the machine learning algorithm sufficiently accurately.

At this point, I'd like to define the people, whom I previously also referred to as "experts" and who I am going to interview, more precisely: I will be interviewing three teachers, who teach at the ArtEz conservatory in Zwolle, who have demonstrable knowledge of the music of J.S.Bach, and generally knowledge of classical music. These people should have provable experience with baroque counterpoint, music theory of the 17th century and generally a good overview of stylistic traits of chorales by J.S.Bach - that's why I have chosen precisely classical music theory, music history and baroque counterpoint teachers from the classical department at ArtEz in Zwolle, as the experts whom I am going to interview. Since those experts will have a more or less similar musical background, for this research, a very detailed interview with 3 persons will be more than satisfying. Also, in order to keep this research possible to accomplish, in a reasonable timeframe, I will only be interviewing teachers from Zwolle. These are the following persons, that will be interviewed:
Ivo van Emmerik;
Harrie Janssen;
Koen van der Meer.

The interview questions can also be found in the appendix and the interview will be held in person. For the interviews I will bring physical print-outs (sheet music) of the generated music, and, in order to aid, I will also offer to listen to a midi-playback using MuseScore on my laptop. The latter is exclusively to help the expert to easier recognise the harmony, the focus will be on the sheet music.

Preceding the interview with the experts, which is the main part of my research, will be two other stages: There will be, first of all, a preparation phase, which is also part of my research: In this preparation stage I will explore literature, in order to understand, how the above mentioned machine learning algorithm (from the paper "DeepBach: a Steerable Model for Bach Chorales Generation"(Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016)") works and also explore literature that depicts chorales by J.S.Bach, in order to be aware of the possible differences between the different machine learning algorithms, later in the interview stage.

Flowing from this preparation stage will be the second stage, the stage of experimentation. In this step, with the help of my, above mentioned, literature, I will decide upon three different parameters for the machine learning algorithm, leading to the generation of three different "fake" chorales.

Data collecting and analysis

During the interview, I will be asking the questions I have outlined in the appendix. Every person will be asked the same questions. During the interview I will be noting, with pen and paper, keywords, that will be mentioned. These keywords (and the amount of times they were mentioned) I will note in a table, in order to have an objective base of comparison. Due to reasons concerning the privacy of the interviewed experts, there will not be voice recordings. However, this is not expected to be a disadvantage for the quality of this research.

3. Results

First, I would like to present the data, which came out of the first and second preparatory stages of my research. In these steps, with the help of above mentioned literature, which I have planned on using, I have found out the, for this research, optimal settings for the machine learning algorithm, which also required to study the literature related to chorales by J.S.Bach (which I also have mentioned above).
For easier usage, I have uploaded the algorithm at the following location: https://github.com/martinkaptein/DeepBachResearch . It is necessary to have the source code be open, so it can be audited. Also, because the software is continuously updated, the results it return may drastically differ. To counteract that, I have uploaded the exact version I have used, together with the usage instructions at the above mentioned address.

Generally speaking chorales by J.S.Bach can be characterised as short exerts that use a four - part harmony (SATB voicing). Historically, they have usually been based on a hymn, that was performed by the soprano, thus harmonised with the Alto, Tenor and Bass voices.[15]

[15] "Bach, J.S. 389 Chorales (Choral-Gesange): SATB (German Language Edition). Kalmus Classic Edition. Alfred Publishing Company, 1985. ISBN 9780769244204. URL https://books.google.fr/books?id= U1-cAAAACAAJ"

These are the following three settings, that, based on my desk research in the first preparatory phase, I have picked to generate three different "fake" J.S.Bach chorales. Based on my desk research from the first preparatory phase, I deemed it necessary to have the algorithm generate three different "fake" J.S.Bach chorales, with different "intensities"[16]. This means that every algorithm, had drastically different settings, with different requirements to computer processing power. I have ranked these three different adjustments of the algorithm from lowest to highest.
This is also necessary for this research, because it provides a more objective base, to be later judged by the committee of experts.

To sum up, each algorithm setting uses different settings:

The first version uses the following settings:

Batch size for training phase: 64
Number of steps per epoch: 250
Number of validation steps: 10
Number of LSTM units: [200, 200]
Size of nun-recurrent hidden layers: 100
Number of Gibbs iterations: 10000
Number of epochs to train the model: 10
Length of unconstrained generations: 100

The second version uses the following settings:

Batch size for training phase: 128
Number of steps per epoch: 500
Number of validation steps: 20
Number of LSTM units: [200, 200]

[16] At this point and in future I will refer to literature, which I have outlined in the introduction and studied during the preparatory first phase of my research.

Size of nun-recurrent hidden layers: 200
Number of Gibbs iterations: 20000
Number of epochs to train the model: 15
Length of unconstrained generations: 100

The third version uses the following settings:

Batch size for training phase: 128
Number of steps per epoch: 500
Number of validation steps: 30
Number of LSTM units: [500, 500]
Size of nun-recurrent hidden layers: 300
Number of Gibbs iterations: 40000
Number of epochs to train the model: 15
Length of unconstrained generations: 100

At this point it is necessary to explain the fact, that precisely above mentioned values have been chosen. As I have mentioned above, it is beneficial to this research to present to each committee of experts three different results of the algorithm, in other words three different, "fake", chorales. To better approximate an answer to my main research question, namely how experts perceive each one of these "fake" chorales in terms of their level of imitation of the composer J.S.Bach, it is more objective to present said experts with a broader base of the complexity of the algorithm, and, thus, as has been deducted by studying the paper "DeepBach: a Steerable Model for Bach Chorales Generation"(Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016)", outputs of the algorithm in the form of chorales. Hence, changes to the algorithm, directly affect the complexity of the "fake" chorale, that is being produced.

To sum up:
The first version uses a very small neural network as depicted by the values " LSTM units", and "hidden layers"). It is also trained for the shortest amount of time on the database of real

J.S.Bach (that is also depicted in my list of literature). Hence, it makes sense to assume, the less complex and musically intricate output in the form of the "fake" chorale.
This will later be cross-referenced, and this thesis will be checked during the 3rd and main part of my research, during three qualitative interviews, with three experts.

The second version, uses the default settings, that are also recommended by the developers of the algorithm, and depicted in above mentioned paper. It could be described as a middle ground between the first and third version of the algorithm.

Finally, the third version of the algorithm uses the highest possible settings for the amount of training iterations and complexity of the neural network, that will "analyse" the training data.

Results of the Interviews

Now, I would like to get to the main part of my research, namely the third phase. In the appendix I will place the exact interview questions I have posed. Also, I will include the outputs of the machine learning algorithm in the form of sheet music.
Moreover, I would like to add the fact, that the output format of the machine learning algorithm was in MIDI-form. To make it more easily readable for the experts, I have transformed the MIDI files into sheet-music using the free software MuseScore in the form of SATB (Soprano, Alto, Tenor, Bass) order.

For this research, it is necessary to explain the data processing process: This I will do in the following paragraphs.

My interview questions can be summarised (and processed) in four blocks:

1. If you didn't know otherwise, would you think this chorale could be composed by J.S.Bach?
2. What are the reasons, according to you, to assume, that this could or couldn't be a real composition by J.S.Bach?

The first two questions have been asked while presenting each expert with each copy of the sheet music separately. For each of the three sheet music copies (=for each of the three

algorithm outputs) I have asked each expert the same questions. As I have argued earlier, the order of the sheet music outputs presented during the interviews was each time the same.

After the expert has seen all three copies and answered the above mentioned two questions, there are the next 2 blocks of questions (summarised):
3. What are the differences between those versions of the machine learning algorithm?
4. Which one of these versions closest matches the style of Bach?

The question blocks 1 and 4 can be answered by a simple answer. Block 1 can be answered in three ways: Yes, No, Maybe. Block 4 can be answered in four ways: Version 1, Version 2, Version 3, neither version.
But how can I quantify which answer each experts has 'selected'?
For the first block of questions I planned to use a simple system: I would just need to note the answer of the expert. For the 4th block I have used a keyword scoring system: I noted with pen and paper the amount of times, each expert exclaimed that version 1, version 2 or version 3, closest matches the style of J.S.Bach, and the one version which received more than 80% of the casted keywords, has been selected as the answer. However, in hindsight, this system was entirely redundant, as for the fourth block of interview question, I also received very clearly interpretable answers.

The remaining blocks 2 and 3 of the interview questions, which in their essence are, by the way, very similar, required a different approach, because these are very open questions. In this case I have just made a list, with a list of (musical) traits, that consisted in each answer, for example a list of differences between each version according to each expert. This is, basically the argumentation for the answers of interview question block 1 and block 4.

In the following I will present the results of my research in the context of above mentioned interview question blocks. For the purpose of presentation, I will make a table.

Interview results block 1:
If you didn't know otherwise, would you think this chorale could be composed by J.S.Bach?

	Algorithm version 1	Algorithm version 2	Algorithm version 3
Respondent 1	No	Yes	No
Respondent 2	No	No	No
Respondent 3	No	No	No

Interview results block 2 (summary):

What are the reasons, according to you, to assume, that this could or couldn't be a real composition by J.S.Bach?

	Algorithm version 1	Algorithm version 2	Algorithm version 3
Respondent 1	No, because: - it is too simple; - it uses a weird melody; - starts and ends in different tonalities (C major to A minor); - bar 2 alto voice weird voice leading.	Yes, because: - Harmonically complex, starts on augmented chord; - Interesting voice leading.	No, because: - no tonic tonality; - primitive counterpoint; - boring melody; - weird and (too) unusual modulation from g minor to C major.
Respondent 2	No, because: - No fermata's (which would be usual for J.S.Bach); - Tone repetition in the alto voice in Bar 2; - Bad melody; - tenor voice too high (register); - bad voice leading.	No, because: - Augmented chord in the beginning; - plagal cadence in the end; - parallel fifth on the second beat in the second Bar; - Bar 6 2nd and 4th beat doesn't resolve.	No, because: - To little chromaticism; - a minor not possible in the context of g minor - Bad C-sharp in Bar 2; - Parallel octave in Bar 4; - No resolution to the tonic on the first beat.
Respondent 3	No, because: - incorrect timing: Generally expected resolution to the tonic on the heavy beat; - It feels like a beginner; - g minor in c major context (in bar 2) is out-of-place; - Impossible bass-line jumps (bar 2 to bar 3); - Parallel fifths in Bar 5; - Bad melody.	No, because: - Bass in Bar 2 not possible; - No fermata's; - No phrases and no phrasing; - End of Bar 3 not possible; - Ending doesn't work/ isn't in the style of J.S.Bach.	No, because: - Tonal context not comprehensible (A7 in g minor); - The ending should not be syncopated, the final chord comes on a weak beat, which is not in the style of Bach. - Parallel octaves in Bar 4; - No phrasing.

Interview results block 3 (summary):

What are the differences between those versions of the machine learning algorithm?

	Perceived differences, based on observations from interview blocks 1 and 2
Respondent 1	- The first version seems to be the most primitive of the three, especially in the context of harmonic complexity and melody. - The second version is much better, especially because of its melody. Out of the three versions it uses the most interesting harmonies. - The third version seems to be a fall-back: It has a very boring melody and much more primitive counter-point.
Respondent 2	- The first version has too many mistakes, especially when compared to the other three versions. - The second version is much better when compared to the first version, and also better when compared to the third version. The latter because the second version uses much more chromatic notes and approaches. - The third version is overall worse than the second version, but it has a slightly superior and better melody.
Respondent 3	- The first version behaves like a beginner, with a bar-to-bar thinking; it does have a lot of mistakes. - The second version is definitely better than the first and third version; it has also mistakes but could be closest comparable to the style of J.S.Bach. - The third version is the worst, especially because it has even more mistakes than the first version, and it is less logical in its harmonies. - Overall, what the three version have in common, is their bar-to-bar approach as opposed to a phrase logic. And this is a trait that is clearly not reminding of J.S.Bach.

Interview results block 4 (summary):

Which one of these versions closest matches the style of Bach?

	Answer
Respondent 1	Version 2
Respondent 2	Version 2
Respondent 3	Neither

Conclusion

The goal of my research was to find out the possibilities of computers, more precisely, a machine learning algorithm in the context of approximating real chorales by J.S.Bach as judged by experts. In other words, the goal of my research is to find out whether J.S.Bach can be imitated using machine learning. It is important to be aware of the capabilities of modern technologies, even in the field of classical music.

The following was my research question: How do experts perceive three different chorales, that have been generated by a machine learning algorithm, with three different adjustments, that has been trained on 389 real chorales by J.S.Bach, in terms of its level of imitation of J.S.Bach?

To answer this question, it was necessary to structure it in the following sub-questions:

1. How is a Bach chorale characterised?
2. How does the machine learning algorithm from the paper "DeepBach: a Steerable Model for Bach Chorales Generation"(Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016) work?
3. How can the machine learning algorithm be adjusted for (three different) results, that approximate chorales by J.S.Bach as close as possible?
4. How do experts perceive differences between a real Bach chorale and three different results of a machine learning algorithm?
5. Which algorithm adjustment yields the best replication of a chorale by J.S.Bach in the perception of experts and why?

To answer this research question I was guided by the above outlined sub-questions, which have been, during the course of this interview, answered.

Now, I would like to get to the conclusion of my research, namely the approximation of an answer to my main research question.

The collected data, in the form of the interviews with the group of experts clearly tends towards a direction: All three results of the machine learning algorithm were clearly seen as not originating from J.S.Bach with numerous flaws. However, all three experts, surprisingly uniformly gravitated towards the opinion, that one particular version, namely the second version of the machine learning algorithm, demonstrated the best result, even with one expert out of the three experts, namely respondent 1, agreeing, that it could be actually a composition by J.S.Bach.

Reasons why respondent 1 thought so, were mentioned in the table overview previously. To summarise, the second version of the machine learning algorithm uses interesting harmonies, thus is harmonically complex and unusual, while also featuring an interesting melody. However, for the majority of the experts this appears to be not enough to convince them, of the possibility, that J.S.Bach might have composed such a chorale. There are just too many mistakes, formal mistakes as well as stylistic mistakes, in the sense that "J.S.Bach just wouldn't write something like that" (according to the pilot group of experts).

What is important to keep in mind, that the experts were already thinking in a very critical way, about the level of imitation of the machine learning algorithm of J.S.Bach, as they already, before the interview, knew from me, that those compositions were in fact not by J.S.Bach. Still, the amount of objectively identifiable mistakes in the three outputs of the machine learning algorithm (which can be traced back from the sheet music) prevailed the reasoning, that the latter were not compositions by J.S.Bach.

Discussion

The above formulated result constitute the positive answering of my research question. However, there are many factors that should be mentioned, which could be different. In the

following I will elaborate on this, and explain, what, according to me, could have been done differently for the purposes of this research.

Still, it is questionable, whether changes to those factors would change the result of my research by a big margin. In the following I will mention the results. But before that, one other thing has to be mentioned: As I have stated in my "onderzoeksontwerp", I could not find any research with the same topic as my research. This hasn't changed a lot, but an interesting thing is that on the internet platform YouTube, there has been done qualitative research, on, almost the same topic as this research (https://www.youtube.com/watch?v=xDqx14lZ_ls). What I find intriguing, that the conclusion is very similar to the conclusion of my research. But this has been published at a point, when I was finished with my pre-research and desk research (this video has been published on March 22 2019). Moreover, this uses a different technology, namely it uses a neural network, which reharmonises a melody, which is submitted by a user.
What I used for my research, DeepBach from "DeepBach: a Steerable Model for Bach Chorales Generation"(Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016)" creates the full chorale on its own using machine learning. There is a fundamental difference.

Now, I would like to elaborate on what could be done differently in this research.

First, my research has been limited to only one source of the algorithm. However, I have sufficiently argued for its sufficient suitability for my research. As the above mentioned YouTube video shows, there has been a lot of development in the topic of machine learning algorithm approximating the music of a composer (it was a system developed by Google). Hence, technology changes and evolves incredibly quickly[17], and there is always a possibility that a much superior system will suddenly appear.

Secondly, there are many possible ways to tweak the machine learning algorithm, which I have used, from the paper (Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016) for different results. However, I have argued to have chosen statistically varied and to my knowledge sufficiently different settings, and yet, settings, which would still make sense to research. Of course, on my computer I have limited computing and processing power, the

[17] see Moore's law

third, most demanding version took about 24 hours to compute. There is quite some possibility, that on a more powerful computer, the results would be very different. Still, that must not necessarily be the case, after all, the design of the software, depicted in the paper (Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016), could be also not fully optimal after all. But that's always the case with technology - it's constantly evolving[18].

What is also important to mention, is the fact, that the group of experts consisted of just three people, which clearly could be perceived as a not sufficiently representative group. Since, the interviews were qualitatively oriented, and, thus, took a very long time to accomplish, due to technical reasons, I was forced to keep the amount of interviews at 3. While I am arguing that, for the purpose of this research, this is fully sufficient, it would be still interesting to interview a larger group of people. Whether this would yield a different conclusion than the conclusion reached by this research, of course, remains a question. Hence, the group of three experts plays the role of a pilot group.

An interesting outcome of my research is the tendency of the group of experts to rank the second version of the machine learning algorithm output as the relatively most likely chorale to be originated from J.S.Bach. Since it is not actually possible to fully comprehend the inner workings of a neural network due to its architecture[19], one can only speculate. The third version took considerably longer to compute - 24 hours as opposed to 5 hours (for the second version), yet, as judged by the committee of experts, from a perspective of a musician and expert, it is far superior. A hypothesis is, that this can be attributed to a phenomena called "overfitting" where the machine learning algorithm, more precisely the neural network, has found a "wrong" approach to imitating J.S.Bach but "thinks" it is "right"[20].

Coming to the end of my research, I would like to think about the implications for society, and, more precisely, for the classical music education facility in ArtEz with this research.
In my opinion, it is a reassuring message, that the works of a genius, that is arguably J.S.Bach cannot be easily "understood" and imitated by a machine learning algorithm.

---

[18] see Moore's law

[19] James, Gareth (2013). An Introduction to Statistical Learning: with Applications in R. Springer. p. 176. ISBN 978-1461471370.

[20] Bishop, C. M. (2006), Pattern Recognition and Machine Learning, Springer, ISBN 978-0-387-31073-2

Still, it would be interesting to better understand the actual creative process of a human being, and why it is so difficult to imitate using machine learning. But this requires a completely different research.

Appendix

Planning

Week 1 and 2: Exploring literature

Week 3 and 4: Conducting experiments

Week 5: Finding experts for interview

Week 6: Conducting interviews

Week 7: Processing interviews

Week 8: Writing research report

Weeks 9-10: Extra margin

Interview questions

- (for each of the three examples I ask (one by one)):

- If you didn't know otherwise, would you think this chorale could be composed by J.S.Bach?
- What is the reason for your answer?
- Are the formal reasons to assume, that this could or couldn't be a real composition by J.S.Bach like:
- harmony (chords)?
- voice-leading?
- rhythms?
- counterpoint?
- speed of harmonic progression?
- cadences?
- and other general stylistic traits? (very open question)

(after we have discussed all three different versions, one by one)

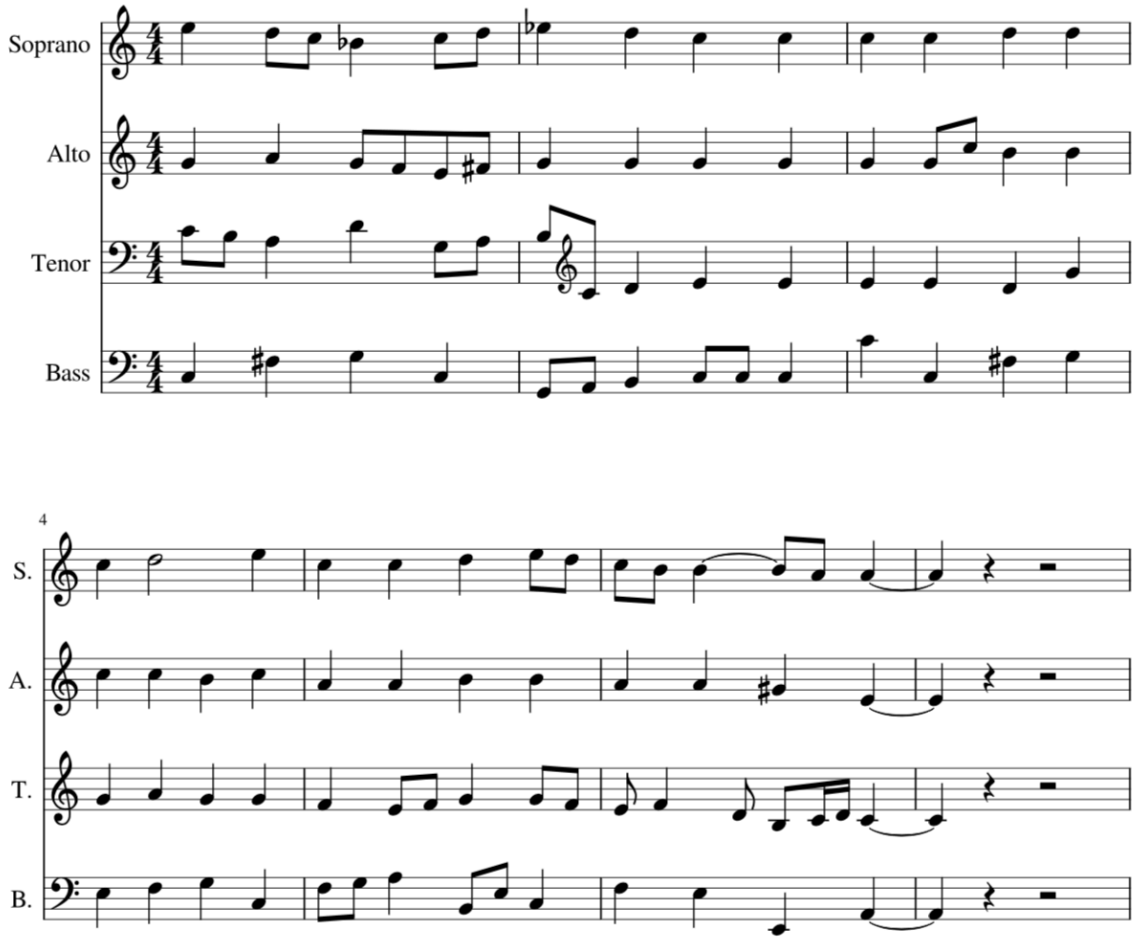
- What are the differences between those versions of the machine learning algorithm?
- Which one of these versions closest matches the style of Bach and why?
- What should be changed (removed, or added) (in the score) so that it would 'sound more like J.S.Bach?

(to close the interview)

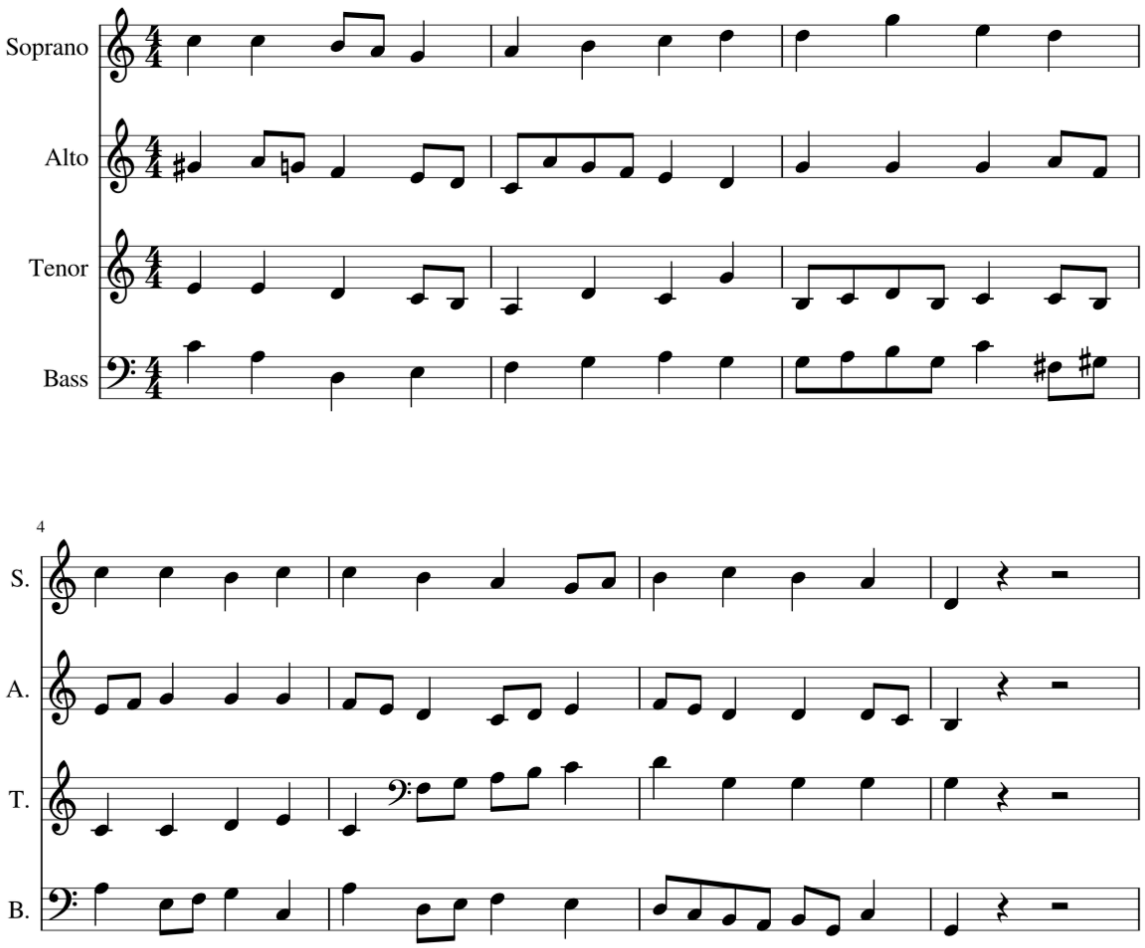
- What do you think about the possibilities of machine learning in regards to its imitation of composers?

Source code of the machine learning algorithm: https://github.com/martinkaptein/DeepBachResearch

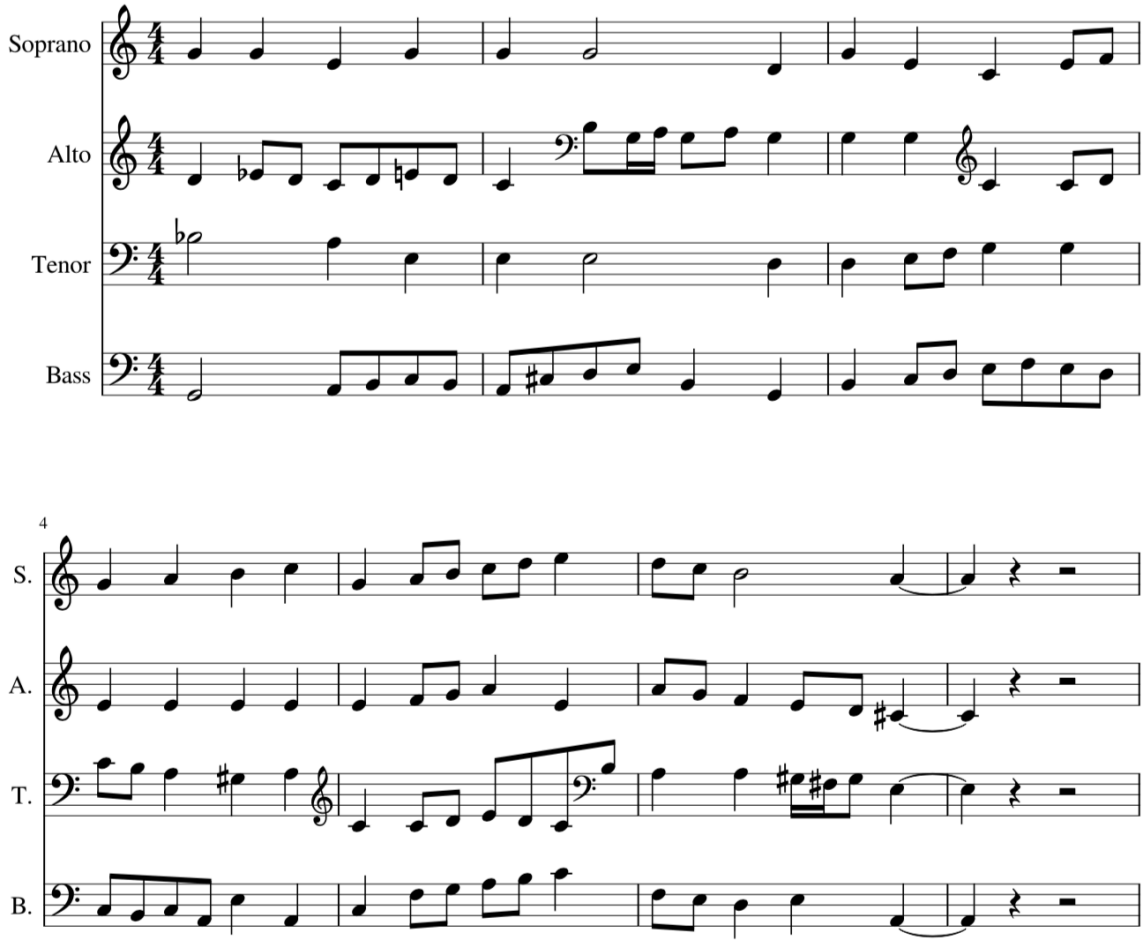
DeepBach research version 1



DeepBach research version 2



DeepBach research version 3



Reference list

Darwin among the Machines". The Press, Christchurch, New Zealand. 13 June 1863.

R. Kohavi and F. Provost, "Glossary of terms," Machine Learning, vol. 30, no. 2–3, pp. 271–274, 1998.

Ebcioglu, Kemal. An expert system for harmonizing fourpart chorales. Computer Music Journal, 12(3):43–51, 1988. ISSN 01489267, 15315169. URL http:// www.jstor.org/stable/3680335. (visited on 20.11.18)

DeepBach: a Steerable Model for Bach Chorales Generation"(Gaëtan Hadjeres, François Pachet, Frank Nielsen, 2016)

James, Gareth (2013). An Introduction to Statistical Learning: with Applications in R. Springer. p. 176. ISBN 978-1461471370.

Liang, Feynman. Bachbot. https://github.com/feynmanliang/bachbot (visited on 29.11.18) , 2016.

Bach, J.S. 389 Chorales (Choral-Gesange): SATB (German Language Edition). Kalmus Classic Edition. Alfred Publishing Company, 1985. ISBN 9780769244204. URL https://books.google.fr/books?id= U1-cAAAACAAJ

Laitz, Steven G. (2008). The Complete Musician (2 ed.). New York, NY: Oxford University Press, Inc. p. 96. ISBN 978-0-19-530108-3.

Bishop, C. M. (2006), Pattern Recognition and Machine Learning, Springer, ISBN 978-0-387-31073-2

Websites:

https://news.sophos.com/en-us/2017/07/24/5-questions-to-ask-about-machine-learning/ (visited on 3.12.18)

https://en.wikipedia.org/wiki/Talos (visited on 3.12.18)

https://emerj.com/ai-glossary-terms/what-is-machine-learning/ (visited on 29.11.18)

https://en.wikipedia.org/wiki/Algorithm (visited on 3.12.18)

https://en.wikipedia.org/wiki/Machine_learning (visited on 3.12.18)

https://en.wikipedia.org/wiki/Artificial_neural_network (visited on 3.12.18)

https://en.wikipedia.org/wiki/Machine_learning (visited on 3.12.18)

http://tones.wolfram.com/about/how-it-works/ (visited on 3.12.18)