

Final Team Project: Music Genre and Composer Classification Using Deep Learning

[Start Assignment](#)

- Due Aug 11 by 11:59pm
- Points 300
- Submitting a text entry box, a website url, a media recording, or a file upload
- Attempts 0
- Allowed Attempts 3

Introduction

Music is a form of art that is ubiquitous and has a rich history. Different composers have created music with their unique styles and compositions. However, identifying the composer of a particular piece of music can be a challenging task, especially for novice musicians or listeners. The proposed project aims to use deep learning techniques to identify the composer of a given piece of music accurately.

Objective

The primary objective of this project is to develop a deep learning model that can predict the composer of a given musical score accurately. The project aims to accomplish this objective by using two deep learning techniques: Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN).

Project Timeline

- Module 2 (by the end of Week 2): The course instructor grouped students into teams of two to three members. Canvas, USD Email, or Slack can be used to find prospective team members.
- Module 4 (by the end of Week 4): Each team's representative submitted the "Team Project Status Update Form."
- Module 7 (by the end of Week 7): Each team should submit deliverables for the course project in the final week:
 1. Project Report
 2. Project Notebook
- It is critical to note that no extensions will be given for any of the final projects' due dates for any reason, and final projects submitted after the final due date will not be graded.

Dataset

The project will use a dataset consisting of musical scores from various composers. [Download the dataset from Kaggle website](https://www.kaggle.com/datasets/blanderbuss/midi-classic-music) [↗ \(https://www.kaggle.com/datasets/blanderbuss/midi-classic-music\)](https://www.kaggle.com/datasets/blanderbuss/midi-classic-music).

The dataset contains the midi files of compositions from well-known classical composers like Bach, Beethoven, Chopin, and Mozart. The dataset should be labeled with the name of the composer for each score. Please only do your prediction only for below composers, therefore you need to select the required composers from the given dataset above.

- 1-Bach
- 2-Beethoven
- 3-Chopin
- 4-Mozart

Methodology

The proposed project will be implemented using the following steps:

1. **Data Collection:** Data is collected and provided to you.
2. **Data Pre-processing:** Convert the musical scores into a format suitable for deep learning models. This involves converting the musical scores into MIDI files and applying data augmentation techniques.
3. **Feature Extraction:** Extract features from the MIDI files, such as notes, chords, and tempo, using music analysis tools.
4. **Model Building:** Develop a deep learning model using LSTM and CNN architectures to classify the musical scores according to the composer.
5. **Model Training:** Train the deep learning model using the pre-processed and feature-extracted data.
6. **Model Evaluation:** Evaluate the performance of the deep learning model using accuracy, precision, and recall metrics.
7. **Model Optimization:** Optimize the deep learning model by fine-tuning hyperparameters.

Deliverables

There are **two** deliverables for this Final Project:

1. **Project Report:** A comprehensive documentation/report that describes the methodology, data pre-processing steps, feature extraction techniques, model architecture, and training process for reproducibility and future reference. Write your technical report in APA 7 style (here is a [Sample Professional Paper \(https://sandiego.instructure.com/courses/19577/files/3232552?wrap=1\)](https://sandiego.instructure.com/courses/19577/files/3232552?wrap=1) format to follow). Please submit the report in PDF format and use the File naming convention DeliverableName-TeamNumber.pdf; for example, **Project_Report-Team1.pdf**
 - Your report should:
 - contain a reference list that includes any external sources, libraries, or frameworks used during the project, including proper citations or acknowledgments.
 - include a concluding section or markdown cell that summarizes the project, highlights key findings, and suggests any potential future improvements or extensions to the work.
2. **Project Notebook:** A Jupyter Notebook file (.ipynb) that contains the entire project code, including data pre-processing, feature extraction, model building, training, evaluation, and any additional analysis or visualizations performed during the project.
 - This deliverable will be exported from a Jupyter Notebook and submitted as a PDF or HTML file.

Conclusion

The proposed project aims to use deep learning techniques to accurately predict the composer of a given musical score. The project will be implemented using LSTM and CNN architectures and will involve data pre-processing,

feature extraction, model building, training, and evaluation. The final model can be used by novice musicians, listeners, and music enthusiasts to identify the composer of a musical piece accurately.

Power Usage for this Project

- You can use Google Colab GPU and TPU in case you need more computation power. Change your runtime in Google Colab notebook to GPU or TPU.
- Another option is to buy the subscription in case you need more computational power (recommended).
 - Please follow this link to do so: [Google Colab Pro+](https://sandiego.instructure.com/courses/19577/files/3232541?wrap=1) (<https://sandiego.instructure.com/courses/19577/files/3232541?wrap=1>).

NOTE: Team members may not get the same grade on the Final Team Project, depending on each team member's level of contribution.

***AI-assisted tools such as ChatGPT, Gemini, and GitHub Copilot should be used to enhance your learning, not replace it. If you use any AI tools in your assignments, you *must* explicitly disclose, cite, and **explain** their contributions (e.g., comments in code, footnotes in reports). Submitting AI-generated code and answers without sufficient understanding, modification, and justification is unacceptable. Proper attribution is essential to maintain academic integrity across all courses and will help position you for future success in this field.

This assignment has [Turnitin](https://help.turnitin.com/integrity/student/canvas/assignments/submitting-an-assignment.htm) [enabled](https://help.turnitin.com/integrity/student/canvas/assignments/submitting-an-assignment.htm) for submissions which means that your instructor will obtain a Similarity Report that identifies specific parts of your writing that may indicate a high level of matching to external content. You are strongly encouraged to review your work without penalty by activating the [Draft Coach extension in your Google Docs](https://help.turnitin.com/integrity/student/draft-coach/using-draft-coach.htm) [prior](https://help.turnitin.com/integrity/student/draft-coach/using-draft-coach.htm) to submitting your work for final grading.

To understand how your work will be assessed, view the scoring rubric below.

Click the **Start Assignment** button above to submit your assignment.

Final Team Project Scoring Rubric

Criteria	Ratings					Pts
Project Report 25%	75 pts Meets or Exceeds Expectations Report thoroughly describes the methodology, data preprocessing steps, feature extraction techniques, model architecture, and training process for reproducibility and future reference.	67.5 pts Approaches Expectations Report generally describes the methodology, data preprocessing steps, feature extraction techniques, model architecture, and training process for reproducibility and future reference.	61.5 pts Below Expectations Report minimally describes the methodology, data preprocessing steps, feature extraction techniques, model architecture, and training process for reproducibility and future reference.	52.5 pts Inadequate Attempt Report is missing one or more of the following: 1. Methodology 2. Data preprocessing steps 3. Feature extraction techniques 4. Model architecture 5. Training process for reproducibility and future reference.	0 pts Non-Performance Non-performance.	75 pts
Project Notebook 65%	195 pts Meets or Exceeds Expectations Team's notebook contains the entire project code in high quality, including all of the following elements: 1. Data preprocessing 2. Feature extraction 3. Model building 4. Training 5. Evaluation 6. Additional analysis or visualizations performed during the project.	175.5 pts Approaches Expectations Team's notebook contains the entire project code, including all of the following elements, but needs minor revision: 1. Data preprocessing 2. Feature extraction 3. Model building 4. Training 5. Evaluation 6. Additional analysis or visualizations performed during the project.	159.9 pts Below Expectations Team's notebook contains the entire project code, including all of the following elements, but needs major revision: 1. Data preprocessing 2. Feature extraction 3. Model building 4. Training 5. Evaluation 6. Additional analysis or visualizations performed during the project.	136.5 pts Inadequate Attempt Team's notebook is missing 1 or more of the following elements: 1. Data preprocessing 2. Feature extraction 3. Model building 4. Training 5. Evaluation 6. Additional analysis or visualizations performed during the project.	0 pts Non-Performance Non-performance.	195 pts
References and	15 pts Meets or Exceeds	13.5 pts Approaches Expectations	12.3 pts Below Expectations	10.5 pts Inadequate Attempt	0 pts Non-Performance	15 pts

Criteria	Ratings					Pts
Citations 5%	Expectations Project includes a references section or markdown cell that lists all external sources, libraries, or frameworks used during the project, including proper citations or acknowledgments in proper APA format.	Project includes a references section or markdown cell that lists all external sources, libraries, or frameworks used during the project, including proper citations or acknowledgments. However, minor revisions are needed in APA format.	Project includes a references section or markdown cell that lists all external sources, libraries, or frameworks used during the project, including proper citations or acknowledgments. However, major revisions are needed in APA format.	Project includes a references section or markdown cell, but does not list all external sources, libraries, or frameworks used during the project, including proper citations or acknowledgments.	Non-performance.	
Conclusion 5%	15 pts Meets or Exceeds Expectations Project includes a thorough concluding section or markdown cell that summarizes the project, highlights key findings, and suggests any potential future improvements or extensions to the work.	13.5 pts Approaches Expectations Project includes an adequate concluding section or markdown cell that summarizes the project, highlights key findings, and suggests any potential future improvements or extensions to the work.	12.3 pts Below Expectations Project includes a vague concluding section or markdown cell that summarizes the project, highlights key findings, and suggests any potential future improvements or extensions to the work.	10.5 pts Inadequate Attempt Project includes a concluding section or markdown cell that is missing one or more of the following elements: 1. Summarizes the project. 2. Highlights key findings. 3. Suggests any potential future improvements or extensions to the work.	0 pts Non-Performance Non-performance.	15 pts
Total Points: 300						