The Genre Factor

Project Presentation - ML Seminar 2023

Henry Krämerkämper
Christopher Breitfeld

13.07.2023

Technische Universität Dortmund

Introduction to the Problem

- Genre classification of songs based on audio features
- Dataset: Thousands of songs with diverse genres
- Challenge: Developing an accurate classification model

Artist	Gorillaz
Url_spotify	https://open.spotify.com/artist/3AA28KZvwAUcZu...
Track	Feel Good Inc.
Album	Demon Days
Album_type	album
Uri	spotify:track:0d28khcov6AiegSCpG5TuT
Danceability	0.818
Energy	0.705
Key	6.0
Loudness	-6.679
Speechiness	0.177
Acousticness	0.00836
Instrumentalness	0.00233
Liveness	0.613
Valence	0.772
Tempo	138.559
Duration_ms	222640.0
Url_youtube	https://www.youtube.com/watch?v=HyHNuVaZJ-k
Title	Gorillaz - Feel Good Inc. (Official Video)
Channel	Gorillaz
Views	693555221.0
Likes	6220896.0
Comments	169907.0
Description	Official HD Video for Gorillaz' fantastic trac...
Licensed	True
official_video	True
Stream	1040234854.0

Figure 1: Example of a song and its attributes.

Description of the Data Set

Dataset from Kaggle: Spotify and YouTube

- Contains statistics of songs on Spotify and YouTube
- including streams on Spotify and number of views on YouTube
- 20.7k entries by 2k artists
- Does NOT include the genre
- Licensed under CC0: Public Domain

Wikidata Query for the Top-Genre of the Artist

- Query artist's Wikidata page for genre names
- Choose a list of broader genres, so that the genres are not too specific
- Match artist to one of the genres on the list, based on the query
- Usage possible under CC-by-SA-3.0

Target of the resulting Dataset

- Target: genre

Network Architecture

- Neural network model for genre classification
- Layers, activation functions, and number of parameters
- Training process and optimization algorithm

Hyperparameter Optimization

- Importance of hyperparameter tuning
- Methods used (e.g., grid search, random search)
- Results of optimization and best hyperparameters

Overtraining Checks

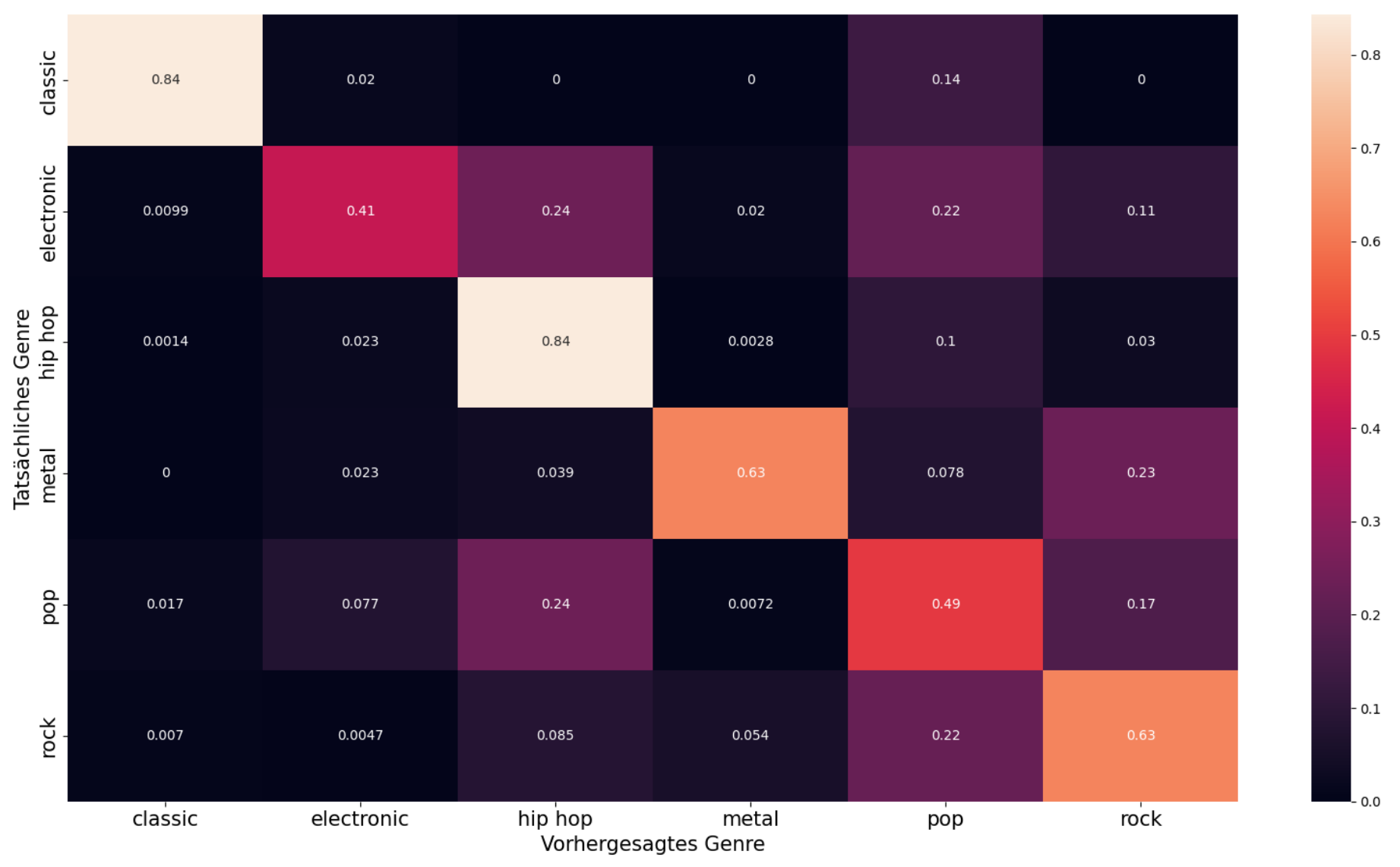
- Preventing overfitting in the model
- Regularization techniques used (e.g., dropout, weight decay)
- Validation and test set performance

Results of our Neural Network

<1>

Results of our Neural Network

<1>



Accuracy and AUC-PR

Alternative Methods

K-nearest-neighbors

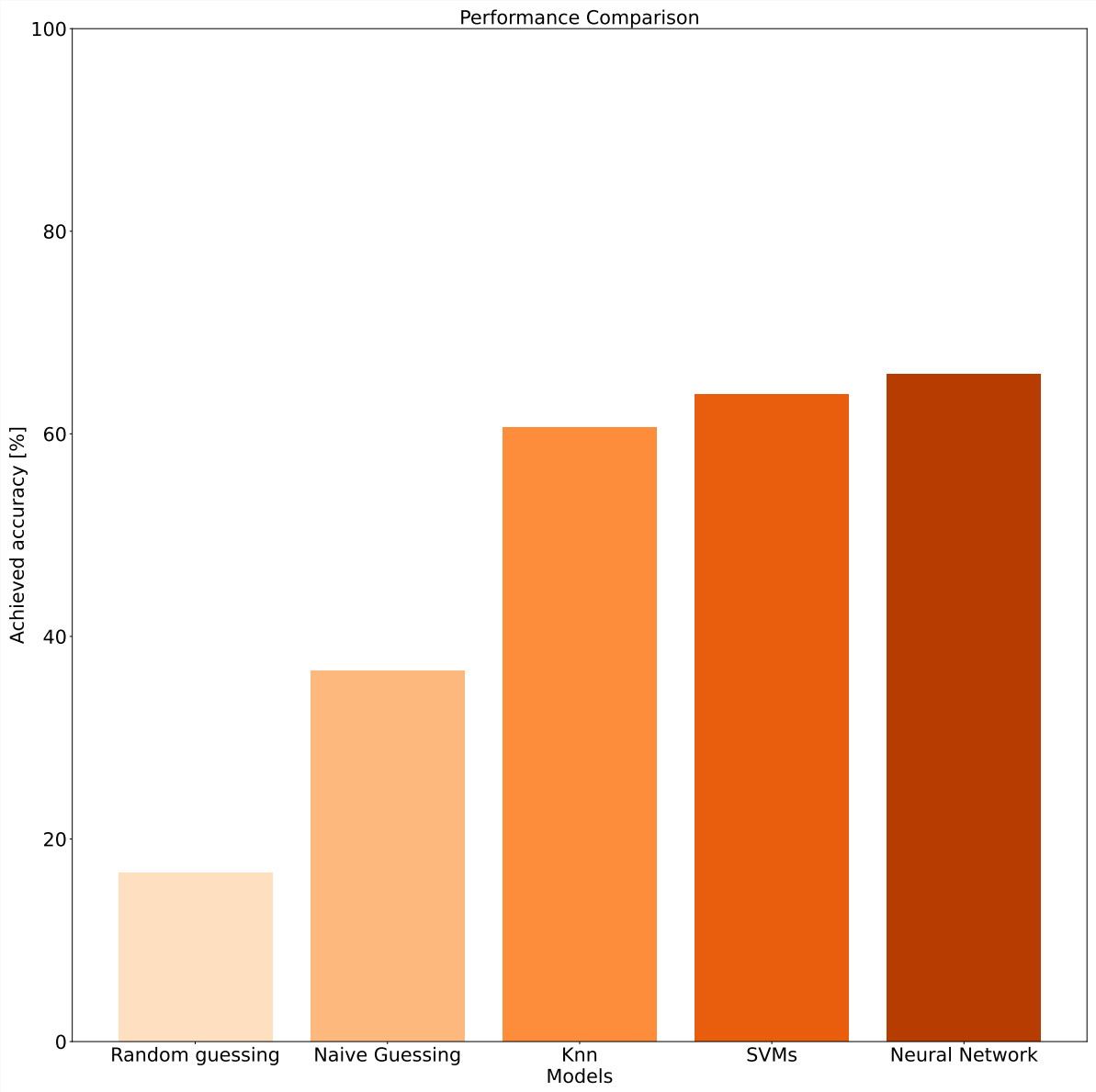
- We use $k = 12$ as it achieves the highest performance
- Results in an accuracy of $60.62\,\%$

Support vector machines

- model that classifies data by finding the hyperplane that maximally separates different categories in a multidimensional space
- We use an One-vs-One approach to be able to do Multiclass-Classification:
  - A separate model is trained for each pair of classes, and a given data point is classified by majority voting among the classifiers
- The used kernel function is the radial basis function (RBF)
- Results in an accuracy of $63.88\,\%$

Conclusions

- NN achieves an accuracy of $65.87\,\%$ on test data.
- Diminishing returns for more complex models, we are constrained by the dataset



Appendix: The features of our dataset

Features

- Track
- Artist
- Url_Spotify
- Album
- Album_type
- Uri
- Danceability
- Energy
- Key
- Loudness
- Speechiness
- Acousticness
- Instrumentalness
- Liveness
- Tempo
- Duration_ms
- Stream
- Url_youtube
- Title
- Channel
- Views
- Likes
- Comments
- Description
- Licensed
- official_video