به­ نام خدا

**دانشکده مهندسی کامپیوتر**

**فرم پیشنهاده­ پروژه ­پاياني، دوره­ی كارشناسی** (ویرایش اردیبهشت 1403)

|  |  |
| --- | --- |
| **عنوان پروژه(فارسی):** | \* طراحی و پیاده‌سازی یک دستیار هوشمند تولید موسیقی بر پایه تحلیل احساسات کاربر |
| **عنوان پروژه (انگلیسی):** | Design and Implementation of an Intelligent Music Generator Based on \* User Emotion Analysis |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | نام دانشجو | شماره دانشجويي | بسته اصلی و فرعی | تعداد واحد گذرانده شده | امضا |
| 1 | محمد امین کیانی | 4003613052 | هوش‌مصنوعی-نرم‌افزار | 130 |  |
| 2 |  |  |  |  |  |

|  |  |
| --- | --- |
| استاد راهنماي پروژه: | امضای استاد راهنما - تاریخ |
| نظر استاد راهنما: |

این پیشنهاده در تاريخ در شوراي گروه مطرح گرديد و

□ بدون تغییر مورد تصويب قرار گرفت.

□ با شرايط زير مورد تصويب قرار گرفت.

□ به دلايل زير مورد تصويب قرار نگرفت.

نام عضو هیئت علمی بررسی­کننده: تاریخ و امضا:

1- موضوع پروژه و اهداف آن را به اختصار شرح دهید.

هدف این پروژه، طراحی و پیاده‌سازی یک سیستم دستیار صوتی هوشمند است که با استفاده از تحلیل احساسات کاربر از ورودی‌های متنی، صوتی و تصویری، اقدام به تولید و یا پیشنهاد موسیقی متناسب با شرایط روحی کاربر می‌نماید. ایده‌ی اصلی این پروژه ایجاد تجربه‌ای شخصی‌سازی‌شده برای کاربران است تا بتوانند در هر موقعیت احساسی، به موسیقی‌ای متناسب و منحصربه‌فرد دسترسی پیدا کنند. این پروژه در مسیر توسعه می‌تواند به‌عنوان پایه‌ای برای اپلیکیشن‌های سلامت روان، موسیقی درمانی، یا توصیه‌گرهای هوشمند نیز مورد استفاده قرار گیرد.

2- روش انجام پروژه را توضيح دهيد.

در این پروژه ابتدا تحلیل احساسات چندرسانه‌ای (متن، صدا و تصویر) با کمک مدل‌های یادگیری عمیق و از طریق کتابخانه‌های مطرح نظیر HuggingFace و TensorFlow انجام خواهد شد. سپس خروجی تحلیل به‌صورت یک برچسب احساسی استخراج شده و به ماژول تولید موسیقی ارسال می‌شود. در نسخه ابتدایی، ممکن است از بانک موسیقی یا MIDIهای آماده استفاده شود، اما در فازهای پیشرفته‌تر، مدل‌های مولد مانند MusicVAE یا Music Transformer جهت ساخت موسیقی سفارشی‌سازی‌شده توسعه می‌یابد. همچنین طراحی رابط کاربری تعاملی با ابزارهایی مانند Gradio جهت دریافت ورودی و پخش خروجی انجام خواهد شد. توجه شود که مدل‌ها و الگوریتم‌های دقیق بسته به شرایط و پیشرفت پروژه ممکن است در طول مسیر تغییر یا بهبود یابند. در انتها، در صورت نیاز یک واسط کاربری ساده برای تست دستیار طراحی می‌شود.

3- آيا اين پروژه و يا مشابه آن قبلاً انجام شده است؟ اگر بله، تفاوت­های اين پروژه با پروژه‌هاي قبلي را (در صورت وجود) ذکر کنید.

بله. پروژه‌هایی در زمینه تحلیل احساسات متنی و نیز تولید موسیقی خودکار با استفاده از هوش مصنوعی قبلاً انجام شده‌اند. با این حال، ترکیب این دو حوزه به‌صورت یکپارچه (end-to-end) و با تمرکز بر کاربرد در دستیارهای صوتی هوشمند، نوآوری کلیدی این پروژه به شمار می‌رود.

در ادامه، به برخی از پروژه‌ها و پژوهش‌های مشابه و تفاوت آن‌ها با این پروژه اشاره می‌شود:

1. MuseNet (OpenAI, 2019): مدل قدرتمندی برای تولید موسیقی چندسبکی با استفاده از Transformer است. با اینکه توانایی بالایی در ساخت موسیقی دارد، اما هیچ‌گونه ارتباطی با تحلیل متن یا احساسات ندارد و صرفاً یک سیستم ژنراتیو موسیقایی است.  
    [(https://openai.com/index/musenet/](%20(OpenAI%20–%20https://openai.com/blog/musenet))
2. AIVA (Artificial Intelligence Virtual Artist): یک آهنگ‌ساز مجازی مبتنی بر هوش مصنوعی است که برای ساخت موسیقی متن فیلم و بازی استفاده می‌شود. ورودی آن پارامترهای موسیقایی است و از تحلیل متن یا احساسات خبری نیست.  
   <( https://www.aiva.ai/>)
3. Amper Music: پلتفرمی برای تولید موسیقی با هوش مصنوعی است که کاربران می‌توانند نوع موسیقی دلخواه را انتخاب کنند. اما این سیستم بر اساس ورودی متنی یا احساسات کار نمی‌کند و بیشتر روی انتخاب سبک و ریتم تمرکز دارد.
4. EmoMTB - Emotion Controlled Music Generation (ArXiv, 2023) : در این مقاله، از توصیفات متنی و صوتی برای کنترل تولید موسیقی احساسی استفاده شده است، اما سیستم end-to-end نبوده و تحلیل مستقیم احساس از متن طبیعی کاربر در لحظه انجام نمی‌شود.
5. Text to Music- A Neural Architecture for Generating Music from Text (NeurIPS Workshop, 2021): این پروژه تلاشی برای ساخت موسیقی از متن توصیفی است، اما تمرکز آن روی کلیدواژه‌هایی مانند "epic", "romantic" و غیره است، نه بر تحلیل عمیق احساسات واقعی کاربران.

تفاوت پروژه حاضر با موارد ذکرشده:

* تحلیل دقیق و لحظه‌ای احساسات متنی کاربران به کمک مدل‌های NLP
* تولید موسیقی خودکار بر پایه احساسات استخراج‌شده، نه صرفاً براساس سبک‌های انتخابی یا کلیدواژه‌ها
* طراحی یک سیستم کاملاً یکپارچه و تعاملی که قابلیت استفاده در دستیارهای صوتی هوشمند مانند Siri یا Alexa را دارد
* تمرکز بر شخصی‌سازی تجربه موسیقایی کاربران در لحظه و با هدف ارتقای تجربه احساسی آن‌ها

این ترکیب منحصربه‌فرد از پردازش زبان طبیعی و تولید موسیقی ژنراتیو، با رویکرد کاربردی و تعامل‌محور، پروژه حاضر را از نمونه‌های قبلی متمایز می‌سازد.

4- طرح تجاری خود دررابطه ‌با تجاری­سازی و به­سودرسانی پروژه­ی انجام­شده را (در صورت وجود) توضیح دهید.

با توجه به علاقه روزافزون کاربران به موسیقی شخصی‌سازی‌شده و کاربردهای سلامت‌محور آن، این سیستم می‌تواند به عنوان پایه‌ای برای اپلیکیشن‌های توصیه‌گر موسیقی، سرویس‌های مدیتیشن یا پلتفرم‌های موسیقی درمانی مورد استفاده قرار گیرد. همچنین نسخه‌های تجاری آن می‌توانند با خدمات اشتراکی یا پرداخت درون‌برنامه‌ای توسعه یابند. در حال حاضر تمرکز پروژه روی نسخه آکادمیک و اولیه سیستم است، ولی پتانسیل تبدیل به محصول تجاری در فازهای بعدی وجود دارد.

کاربردهای احتمالی:

* ساخت اپلیکیشنی برای ساخت موسیقی شخصی‌سازی‌شده برای حال روحی کاربر
* استفاده در بازی‌ها برای تولید موسیقی دینامیک بسته به شرایط داستانی یا تصمیمات بازیکن
* استفاده در مراقبت سلامت روان و مدیتیشن با تولید موسیقی آرامش‌بخش بر اساس احساسات کاربران

5- امكانات موردنیاز جهت انجام پروژه را ذكر نماييد.

* لپ‌تاپ یا سیستم با کارت گرافیک مناسب (حداقل 8 GB RAM و GPU برای آموزش مدل‌ها)
* محیط‌های برنامه‌نویسی Python، Jupyter Notebook، Google Colab
* دسترسی به کتابخانه‌های Python مانند transformers ، tensorflow، pydub، gradio، librosa
* دیتاست‌های مربوط به احساسات متنی و نت‌های موسیقی (MIDI)
* هدفون و ابزارهای شنیداری برای تست کیفیت موسیقی تولیدشده

6- مهم‌ترین منابع و مراجع لازم برای انجام پروژه­ی پيشنهادي را نام ببرید.

این منابع شامل مقالات مرجع، پروژه‌های تحقیقاتی، دیتاست‌های استاندارد و کتابخانه‌های پایتون هستند که نقش اساسی در پیاده‌سازی پروژه خواهند داشت. لازم به ذکر است که ابزارها، منابع و دیتاست‌های ذکرشده در این بخش، بر اساس بررسی‌های اولیه انتخاب شده‌اند و ممکن است در طول انجام پروژه با توجه به محدودیت‌های فنی، به‌روزرسانی‌های علمی یا تجربیات عملی، جایگزین یا به‌روزرسانی شوند. بنابراین وابستگی کامل و قطعی به این موارد وجود ندارد و مسیر توسعه‌ی پروژه می‌تواند انعطاف‌پذیر و قابل تطبیق با شرایط واقعی باشد.

* منابع داده (Datasets):
  + GoEmotions –Google Research از  
    شامل 58000جمله با برچسب 27 احساس مختلف + حالت خنثی
  + PerSent (Persian Sentiment Dataset)

دیتاست احساسی فارسی برای تحلیل احساسات متنی

* مقالات و پروژه‌های علمی پایه:
  + Devlin, J. et al. (2019).  
    **BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding**  
    مرجع اصلی در حوزه مدل‌های زبانی برای درک عمیق متن
  + Huang, C. A. et al. (2018).  
    **Music Transformer: Generating Music with Long-Term Structure**  
    معرفی یکی از بهترین مدل‌های تولید موسیقی بر پایه Transformer
  + Ghosal, D. et al. (2021).  
    **Text to Music: A Neural Architecture for Generating Music from Text**  
    NeurIPS ارائه معماری عصبی برای تولید موسیقی از متن در کارگاه
  + Yoo, J. et al. (2023).  
    **EmoMTB: Emotion Controlled Music Generation using Text and Audio Descriptions**مقاله‌ای در زمینه کنترل موسیقی احساسی با توصیف متنی و صوتی
* ابزارها و کتابخانه‌های کلیدی:
  + Google Magenta Project

پروژه‌ای متن‌باز شامل مدل‌های تولید موسیقی و ابزارهای صوتی

* + HuggingFace Transformers

کتابخانه‌ای برای استفاده از مدل‌های پردازش زبان

* + Music21

کتابخانه‌ای برای تحلیل موسیقی و تبدیل ساختارهای نت‌نویسی به داده دیجیتال

* + PyDub

کتابخانه‌ای برای پردازش و ویرایش فایل‌های صوتی در پایتون (تبدیل، افکت‌گذاری، برش)