



فرم پیشنهاد پروژه پایانی، دوره‌ی کارشناسی (ویرایش اردیبهشت ۱۴۰۳)

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 دانشجویی	بسته اصلی و فرعی	تعداد واحد گذرانده شده	امضا
محمد امین کیانی	۴۰۰۳۶۱۳۰۵۲	هوش مصنوعی - نرم افزار	۱۳۰	
۲				

	استاد راهنمای پروژه:
	نظر استاد راهنما:
امضای استاد راهنما - تاریخ	

این پیشنهاد در تاریخ

در شورای گروه

مطرح گردید و

☐ بدون تغییر مورد تصویب قرار گرفت.

☐ با شرایط زیر مورد تصویب قرار گرفت.

☐ به دلایل زیر مورد تصویب قرار نگرفت.

نام عضو هیئت علمی بررسی کننده:

تاریخ و امضا:

۱- 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 منحصر به فرد دسترسی پیدا کنند. این پروژه در مسیر توسعه می‌تواند به عنوان پایه‌ای برای اپلیکیشن‌های سلامت روان، موسیقی درمانی، یا توصیه‌گرهای هوشمند نیز مورد استفاده قرار گیرد.

۲- روش انجام پروژه را توضیح دهید.

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 به صورت یک برجسب احساسی استخراج شده و به ماژول تولید موسیقی ارسال می‌شود. در نسخه ابتدایی، ممکن است از بانک موسیقی یا MIDI های 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.

۳- آیا این پروژه و یا مشابه آن قبلاً انجام شده است؟ اگر بله، تفاوت‌های این پروژه با پروژه‌های قبلی 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 به صورت 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:

۱. 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/>

۲. 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/>

۳. 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.

۴. 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.

۵. 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 ذکر شده:

- 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 منحصر به فرد از پردازش زبان طبیعی و تولید موسیقی ژنراتیو، با رویکرد کاربردی و تعامل محور، پروژه حاضر را از نمونه‌های قبلی متمایز می‌سازد.

۴- طرح تجاری خود در رابطه با تجاری سازی و به سودرسانی پروژه ای انجام شده را (در صورت وجود) توضیح دهید.

با توجه به علاقه روزافزون کاربران به موسیقی شخصی سازی شده و کاربردهای سلامت محور آن، این سیستم می تواند به عنوان پایه ای برای اپلیکیشن های توصیه گر موسیقی، سرویس های مدیتیشن یا پلتفرم های موسیقی درمانی مورد استفاده قرار گیرد. همچنین نسخه های تجاری آن می توانند با خدمات اشتراکی یا پرداخت درون برنامه ای 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 شخصی سازی شده برای حال روحی کاربر
- استفاده در بازی ها برای تولید موسیقی دینامیک بسته به شرایط داستانی یا تصمیمات بازیکن
- استفاده در مراقبت سلامت روان و مدیتیشن با تولید موسیقی آرامش بخش بر اساس احساسات کاربران

۵- امکانات مورد نیاز جهت انجام پروژه را ذکر نمایید.

- لپ تاپ یا سیستم با کارت گرافیک مناسب (حداقل 8 GB RAM و GPU برای آموزش مدل ها)
- محیط های برنامه نویسی Python ، Jupyter Notebook ، Google Colab
- دسترسی به کتابخانه های Python مانند librosa ، gradio ، pydub ، tensorflow ، transformers
- دیتاست های مربوط به احساسات متنی و نت های موسیقی (MIDI)
- هدفون و ابزارهای شنیداری برای تست کیفیت موسیقی تولید شده

۶- مهم ترین منابع و مراجع لازم برای انجام پروژه ای پیشنهادی را نام ببرید.

این منابع شامل مقالات مرجع، پروژه های تحقیقاتی، دیتاست های استاندارد و کتابخانه های پایتون هستند که نقش اساسی در پیاده سازی پروژه خواهند داشت. لازم به ذکر است که ابزارها، منابع و دیتاست های ذکر شده در این بخش، بر اساس بررسی های اولیه انتخاب شده اند و ممکن است در طول انجام پروژه با توجه به محدودیت های فنی، به روزرسانی های علمی یا تجربیات عملی، جایگزین یا به روزرسانی شوند. بنابراین وابستگی کامل و قطعی به این موارد وجود ندارد و مسیر توسعه ی پروژه می تواند انعطاف پذیر و قابل تطبیق با شرایط واقعی باشد.

• منابع داده (Datasets):

- GoEmotions – Google Research از شامل ۵۸۰۰۰ جمله با برچسب ۲۷ احساس مختلف + حالت خنثی
- PerSent (Persian Sentiment Dataset) دیتاست احساسی فارسی برای تحلیل احساسات متنی

• مقالات و پروژه های علمی پایه:

- Devlin, J. et al. (2019).  
BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding  
مرجع اصلی در حوزه مدل های زبانی برای درک عمیق متن
- Huang, C. A. et al. (2018).  
Music Transformer: Generating Music with Long-Term Structure  
Transformer معرفی یکی از بهترین مدل های تولید موسیقی بر پایه
- Ghosal, D. et al. (2021).  
Text to Music: A Neural Architecture for Generating Music from Text  
ارائه معماری عصبی برای تولید موسیقی از متن در کارگاه NeurIPS
- Yoo, J. et al. (2023).  
EmoMTB: Emotion Controlled Music Generation using Text and Audio Descriptions  
مقاله ای 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)