



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

عنوان پروژه (فارسی): * تولید موسیقی مبتنی بر تحلیل احساسات

* Emotion-Aware Music Generation

عنوان پروژه (انگلیسی):

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

مطرح گردید و

در شورای گروه

این پیشنهاد در تاریخ

☐ بدون تغییر مورد تصویب قرار گرفت.

☐ با شرایط زیر مورد تصویب قرار گرفت.

☐ به دلایل زیر مورد تصویب قرار نگرفت.

تاریخ و امضا:

نام عضو هیئت علمی بررسی کننده:

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

هدف این پروژه طراحی و پیاده‌سازی یک سیستم هوشمند است که ابتدا با استفاده از تکنیک‌های پیشرفته یادگیری عمیق و پردازش زبان طبیعی (NLP)، احساسات موجود در متون کاربران را تحلیل کرده و در مرحله بعد، موسیقی‌ای متناسب با حال و هوای آن احساس تولید کند. این سیستم می‌تواند در بستر یک دستیار صوتی هوشمند به کار گرفته شود تا بدون نیاز به آهنگ‌ساز انسانی، موسیقی‌های غنی و احساسی در سبک‌های مختلف خلق کند.

اهداف اصلی پروژه:

- پیاده‌سازی ماژول تحلیل احساسات متنی با دقت بالا
- طراحی ماژول تولید موسیقی احساسی با استفاده از مدل‌های ژنراتیو (مانند Transformer، VAE یا Diffusion Model)
- انطباق دو ماژول جهت خلق تجربه‌ای طبیعی، دلنشین و خودکار در تعامل انسان و ماشین

۲- روش انجام پروژه را توضیح دهید.

پروژه در دو فاز اصلی انجام خواهد شد:

فاز اول: تحلیل احساسات متنی

- استفاده از مدل‌های پیش‌آموزش‌دیده مانند BERT، RoBERTa یا DistilBERT برای استخراج ویژگی‌های متنی و تحلیل احساسات (مثبت، منفی، خنثی و نیز حالات پیچیده‌تری مثل غم، شادی، اضطراب، هیجان).
- آموزش مدل بر اساس دیتاست‌های معتبر مانند EmotionLines، GoEmotions یا datasets فارسی مانند PerSent یا Emotion-Farsi.

فاز دوم: تولید موسیقی متناسب با احساس

- استفاده از مدل‌های تولیدی مانند Music Transformer یا RNN-VAE برای ساخت ملودی و هارمونی متناسب با احساسات تشخیص داده‌شده.
 - تبدیل احساسات تحلیل‌شده به پارامترهای موسیقی مانند سرعت (BPM)، گام (Scale)، ابزار موسیقی (Instrument) و ریتم.
 - در صورت نیاز، تولید MIDI و تبدیل آن به فایل صوتی نهایی WAV یا MP3 با استفاده از ابزارهایی مانند FluidSynth یا Magenta.
- 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 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:

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. A Neural Architecture for Generating Music from Text (NeurIPS Workshop, 2021) -Text to Music :

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 مانند BERT یا RoBERTa
 - 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.

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

این پروژه پتانسیل بالایی برای تجاری‌سازی دارد، به‌ویژه در حوزه اپلیکیشن‌های موبایل، بازی‌سازی، مدیتیشن، مراقبت روانی و تعامل انسان-ماشین. 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

۵- امکانات موردنیاز جهت انجام پروژه را ذکر نمایید.

- لپ‌تاپ یا سیستم با کارت گرافیک مناسب (حداقل 8 GB RAM و GPU برای آموزش مدل‌ها)
- محیط‌های برنامه‌نویسی Python، Google Colab، Jupyter Notebook
- کتابخانه‌های ML و NLP مانند TensorFlow، PyTorch، Transformers (HuggingFace)، Magenta، Music21
- دیتاست‌های مربوط به احساسات متنی و نت‌های موسیقی (MIDI)
- هدفون و ابزارهای شنیداری برای تست کیفیت موسیقی تولیدشده

۶- مهم‌ترین منابع و مراجع لازم برای انجام پروژه‌ی پیشنهادی را نام ببرید.

این منابع شامل مقالات مرجع، پروژه‌های تحقیقاتی، دیتاست‌های استاندارد و کتابخانه‌های پایتون هستند که نقش اساسی در پیاده‌سازی پروژه خواهند داشت.

• 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 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
Transformer معرفی یکی از بهترین مدل‌های 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
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)