به نام خدا

گزارش کار پروژه مورخ 29/03/1400

محمدمهدی برقی

* **مطالعه مقالات مورد بررسی برای پیدا کردن ویژگی های دیتاست لازم**

طی بررسی های شکل گرفته مقالات مرتبط با پروژه مطالعه و بررسی شدند و تحلیل دقیق تر مقالات مرتبط با مدل انتخابی برخی ویژگی های دیتاست مورد نیاز ارزیابی شدند. تعدادی از مقالات مورد بررسی به شرح زیر اند:

* Text to Speech Synthesis

در این مقاله به شرحی کلی از سیر و روند تکنولوژی های مورد استفاده در تبدیل متن به صدای طبیعی انسان پرداخته می شود و ضمن تحلیل دقیق برخی الگوریتم های این کار به توصیف کلی فرایند تبدیل می پردازد

* Deep Voice
* Deep Voice2
* Deep Voice3

این مقالات به بررسی مدل Deep Voice می پردازند، مدلی قدرتمند که از سال 2018 بررسی های پژوهش های بسیاری بر روی آن انجام شده و در نهایت پژوهشگران به نسل فعلی این مدل یعنی Deep Voice3 دست یافتند که توانسته تاکنون بازده و عملکرد مناسبی را از خود نشان دهد و صدایی نسبتا طبیعی اما همچنان دارای نیوز و پارازیت های غیر طبیعی را تولید کند

* Tacotron
* Tacotron2

مدل Tacotron نیز از مدل های هوش مصنوعی پیشرفته ای برای تبدیل متن به صدا هست که با توجه به کیفیت بهتر و نویز کمتر آن صدای طبیعی تری را در نهایت ارائه می دهد.

از طرف دیگر این مدل می تواند عملکرد مناسب تری را برای زبان های مثل فارسی و عربی که دارای اعراب و واج ها هستند داشته باشد بنابراین با بررسی های صورت گرفته در نهایت این مدل به عنوان مدل اصلی پروژه مورد پژوهش و توسعه قرار گرفت.

* WaveNet

از جمله مدل هایی هست که برای تبدیل نهایی طول موج صوتی تولید شده در سایر مدل ها به صوت مورد استفاده قرار می گیرد و از آن می توان برای تبدیل امواج صوتی تولیدی به صوت استفاده کرد

* **بررسی دیتاست های موجود برای استفاده :  
  در طی ادامه تحقیقات و بررسی ها دیتاست های مختلفی مورد ارزیابی و بررسی قرار گرفتند که در ادامه به آنها اشاره می شود.**
* **دیتاست دکتر زینالی:**

**این دیتاست شامل 100 ساعت دیتاست با انواع کیفیت و به صورت چند گوینده است و برای پروژه تبدیل صوت به متن مورد استفاده قرار گرفته است.**

* **دیتاست فارسی دات بزرگ:**

**حدود هشتاد ساعت دیتاست با کیفیت استدیویی تهیه شده توسط پژوهشگاه توسعه فناوری های پیشرفته**

* **دیتاست موزیلا:**

**شامل 321 ساعت فایل صوتی فارسی با کیفیت های مختلف و به صورت چند گویندگی**

* **کتاب های صوتی:**

**بررسی سه کتاب صوتی :   
با خانمان، طلای درون، عطر خوش مرگ با صدای احسان کرمی که متن و صوت هر سه کتاب ها موجود بود و دیتاستی حدود 38 ساعت به صورت تک گویندگی داشتیم.**

* **مشارکت در تهیه دیتاست ( حدود چهار ساعت از دیتاست)**

**در نهایت دیتاست موزیلا انتخاب شد و تیم دیتاست برای انجام خواسته تیم نرم افزار و تبدیل دیتاست های صوتی به متن فارسی و فینگلیش وارد عمل شدند و من نیز در تهیه حدود 4 ساعت از ده ساعت تهیه شده همکاری داشتم**

* **تهیه لیستی 30000 کلمه از کلمات فارسی فینگلیش**

**در طی فرایند تبدیل فایل های صوتی به متن فینگلیش در نهایت لغت نامه ای شامل 30000 کلمه فارسی و به معادل آن کلمه ی فینگلیش تهیه و تنظیم شد.**

* **برنامه پایتون برای تبدیل فارسی به فینگلیش**

**و در نهایت برنامه ای برای اتوماتیک سازی فرایند تبدیل متن فارسی به فینگلیش و بالا بردن سرعت روند تهیه دیتاست به وجود آمد.**

* **تحلیل سایت همراه اول برای استخراج کلمات کلیدی مرتبط با پروژه**

**در طی جلسات مشترک با تیم همراه اول به این نتیجه رسیدیم که مدل نهایی در کاربرد خاصی مورد استفاده قرار خواهد گرفت بنابراین با تحلیل و بررسی بخش های مختلف سایت همراه اول و استخراج کلمات کلیدی مرتبط با آنها شروع به ایجاد دیتاستی مربوط با پروژه و کاربرد اصلی کردیم.**