



دانشگاه صنعتی امیرکبیر  
(پلی تکنیک تهران)  
دانشکده مهندسی کامپیوتر

گزارش کارآموزی

محل کارآموزی  
شرکت راهکار پردازش ژرف (روشن)

استاد کارآموزی  
دکتر مریم امیرمزلقانی

کارآموز  
مهدی نیک نژاد - ۹۶۳۰۰۳۹

تابستان ۱۴۰۰

## سپاسگزاری

اینجانب مهدی نیک نژاد مراتب امتنان خود را نسبت به استاد کارآموزی و همکاران عزیزم در شرکت روشن که در گذراندن این دوره کارآموزی همواره مرا یاری نموده‌اند، ابراز می‌دارم.

مهدی نیک نژاد  
شهریور ۱۴۰۰

## چکیده

در دوره کارآموزی که در شرکت راهکار پردازش ژرف (شرکت روشن) گذرانده شد، هدف اصلی یادگیری مفاهیمی در حوزه هوش مصنوعی بود. در این راه، ابتدا پروژه‌ای آزمایشی تحت عنوان «نرمال سازی زبان محاوره با استفاده از ابزار هضم» در حوزه پردازش زبان طبیعی پیاده سازی شد تا پیش‌نیازهایی مانند یادگیری کتابخانه‌های Hazm و Tensorflow برطرف شوند و مقدمات ورود به فضای واقعی پیاده سازی یک سامانه فراهم شود. پس از تأیید توانمندی‌های کارآموز، پروژه جدیدی آغاز شد که در حوزه پردازش تصویر بود و تجربه‌های بسیاری در این راه کسب شد. سامانه‌ای که به عهده من گذاشته شد، «ابزار الفبا» از محصولات شرکت بود و می‌بایست داده‌های قطعه بندی برای سند هویتی کارت ملی تولید می‌کردم.

## واژه‌های کلیدی:

پردازش زبان طبیعی، نرمال‌سازی، ترجمه ماشین، پردازش تصویر، شبکه عصبی و segmentation

# فهرست مطالب

صفحه

عنوان

۱	.....	مقدمه
۲	.....	۱-۱ مقدمه
۴	.....	۲ معرفی محل کارآموزی
۵	.....	۱-۲ تاریخچه
۵	.....	۲-۲ معرفی کلی شرکت
۵	.....	۳-۲ ویژگی‌های ابزار الفبا
۵	.....	۱-۳-۲ خواندن متن با قلم‌ها و اندازه‌های مختلف
۶	.....	۲-۳-۲ خواندن نویسه‌های فارسی و انگلیسی
۶	.....	۳-۳-۲ تحلیل ساختار سند
۶	.....	۴-۳-۲ اصلاح چرخش و نویز تصویر در سند
۷	.....	۴-۲ اطلاعات تماس شرکت راهکار پردازش ژرف
۹	.....	۳ فعالیت‌ها و تجربیات کارآموزی
۱۰	.....	۱-۳ شرح موضوع کارآموزی
۱۰	.....	۱-۱-۳ پروژه نرمال‌سازی زبان محاوره
۱۰	.....	۱-۱-۱-۳ شبکه عصبی بازگشتی
۱۱	.....	۲-۱-۱-۳ مدل زبانی
۱۱	.....	۳-۱-۱-۳ تاریخچه ترجمه ماشین
۱۱	.....	۴-۱-۱-۳ ترجمه ماشینی عصبی
۱۲	.....	۵-۱-۱-۳ مدل seq2seq
۱۳	.....	۶-۱-۱-۳ مکانیزم توجه (attention)
۱۳	.....	۲-۱-۳ پروژه تولید داده‌های قطعه‌بندی
۱۴	.....	۴ جمع‌بندی و نتیجه‌گیری
۱۵	.....	۱-۴ نتیجه‌گیری
۱۵	.....	۲-۴ کارهای آینده
۱۶	.....	منابع و مراجع
۱۷	.....	پیوست
۱۷	.....	گزارش‌های کارآموزی
۱۸	.....	گزارش شماره ۱

۲۰	.....	گزارش شماره ۲
۲۱	.....	گزارش شماره ۳
۲۳	.....	گزارش شماره ۴
۲۷	.....	گزارش شماره ۵
۲۸	.....	گزارش شماره ۶
۳۰	.....	گزارش شماره ۷
۳۱	.....	گزارش شماره ۸
۳۳	.....	گزارش شماره ۹

# فهرست اشکال

شکل

صفحه

۵	۱-۲	قلم‌های مختلف
۶	۲-۲	متون فارسی و انگلیسی
۶	۳-۲	ساختار سند
۷	۴-۲	اصلاح چرخش سند
۱۰	۱-۳	شبکه عصبی بازگشتی
۱۲	۲-۳	مدل sequence-to-sequence
۲۳	۱	مجموعه دادگان تسک اولیه
۲۵	۲	کدگذاری حریصانه
۲۵	۳	کدگذاری جستجوی پرتو محلی
۳۰	۴	نمایش خروجی مدل
۳۱	۵	نمایش خروجی مدل بهبودیافته
۳۳	۶	دسترسی ریشه
۳۴	۷	اتصال به vpn
۳۴	۸	سند هویتی (کارت ملی) تولیدشده با کد

## فهرست جداول

صفحه

جدول

۲۴ . . . . .	۱ مقایسه ابزار هضم و ابزار فارسیار با متن اصلی
--------------	--

# فصل اول

## مقدمه



## ۱-۱ مقدمه

داده‌ها بسیار زیاد، گوناگون و با ارزش شده‌اند. اهیت و ارزش داده‌های زیاد مربوط می‌شود به معانی و پاسخ‌هایی که میان آنها نهفته است.

در کسب و کارها، نرم‌افزارهای زیادی با کتاب، مقاله، روزنامه، مجله، فاکتور، صورتحساب، فکس و به طور خلاصه کاغذ کار می‌کنند که محتوای آنها به هر دلیلی با pdf سروکار دارد. این اسناد، اگرچه به ظاهر حروف چینی شده‌اند، اما به یک ابزار مانند الفبا نیاز دارند.

در سازمان، انبوه پرونده‌ها و نامه‌ها در اتاق‌های بایگانی خاک می‌خورند و منتظر نویسه‌خوان فارسی هستند. عنوان «دولت الکترونیک» که رسانه‌ها راجع به آن زیاد گفته‌اند و شنیده‌ایم، چنین پیشنیازی دارد و قبل از وقوع آن بی‌گمان باید سالها پرونده غیرالکترونیک، نویسه‌خوانی شود.

کتابخانه‌های پیشگام، در سالهای اخیر به سراغ اسکن<sup>۱</sup> کردن صفحات کتاب‌ها رفته‌اند. این حجم عظیم از صفحات اسکن شده که تاکنون تنها جنبه نگداری کتاب‌ها در قالب الکترونیکی را داشته‌اند، با استفاده از ابزار نویسه‌خوان می‌توانند قابل جستجو و بازیابی شوند.

تردید نیست که دانشگاه‌ها همواره منبع استعدادهای این کشور بوده و هستند. به همین خاطر شرکت‌های داخلی به دنبال دانشجویان مستعد و فعالی هستند که بتوانند مشکلات خود در زمینه‌ی ارائه راهکارهای مناسب برای حرکت به سمت دولت الکترونیک در زمان کم و با بهترین کیفیت حل نمایند. دوره‌های کارآموزی فرصت مناسبی برای این شرکتها است تا این جوانان را رصد کرده و نیروهای خود را از همین طریق جذب کنند.

شرکت راهکار پردازش ژرف که به عنوان شرکت روشن شناخته می‌شود، نیز از جمله شرکت‌هایی است که در جهت اهداف عنوان شده قدم برمی‌دارد. در این دوره کارآموزی نیز در راستای پیش‌برد همین هدف گام برداشته شد.

در ادامه‌ی این گزارش طی یک فصل نگاهی خواهیم داشت به شرکت روشن که در آن دوره کارآموزی خود را گذراندم. پس از آن، به شرح مفصل پروژه‌هایی که در مدت کارآموزی به آنها مشغول بودم و همچنین فعالیتهایی که در این راستا و یا در طول انجام این مهم انجام داده‌ام می‌پردازم. با گذر از این مراحل، بخش جمع‌بندی پایان این مسیر گزارش خواهد بود.

---

<sup>۱</sup>Scan



## فصل دوم

### معرفی محل کارآموزی

## ۱-۲ تاریخچه

شرکت راهکار پردازش ژرف، با هسته اصلی از نخبگان دانشگاه علم و صنعت ایران فعالیت خود را از سال ۱۳۹۵ به طور رسمی آغاز نمود و در ادامه توانست از نیروهای متخصص و متعهد دانشگاه های طراز اول کشور بهره گیرد.

## ۲-۲ معرفی کلی شرکت

راهکار پردازش ژرف با استفاده از روزآمدترین فناوری یادگیری ماشین، تحت عنوان یادگیری عمیق محصولات متعددی را در حوزه پردازش تصویر، متن و گفتار ارائه نموده است که از بین آنها دو محصول «الفبا» و «کشف» به عنوان محصول دانش بنیان نوع ۱ مورد تایید قرار گرفته اند. هدف اصلی شرکت، ارائه راهکارهای نرم افزاری مبتنی بر هوش مصنوعی جهت پاسخگویی به نیازهای کشور و متناسب با فرهنگ ایرانی و اسلامی می باشد و در چشم انداز خود در نظر دارد محصولاتی را در حوزه یادگیری عمیق و در کلاس جهانی عرضه نماید.

## ۳-۲ ویژگی های ابزار الفبا

نویسه خوان ابزاری است که نوشته ها را از دل تصاویر متنی بیرون می کشد و کاغذ را وارد دنیای رایانه می کند. در دهه اخیر، چند ابزار نویسه خوان فارسی آماده شده اند؛ اما کیفیت آنها نیازهای مخاطب را به خوبی برطرف نکرده است. حالا گروه روشن به عنوان زیر مجموعه ای از شرکت راهکار پردازش ژرف با استفاده از فناوری روز دنیا، ابزار «الفبا» را آماده کرده است تا اسناد را سریع و دقیق تحلیل کند.

## ۱-۳-۲ خواندن متن با قلم ها و اندازه های مختلف

در صفحات متن معمولاً عبارات بر اساس اهمیت شان با قلم و اندازه متفاوت نوشته می شوند. الفبا صفحه را می خواند در حالیکه چند قلم با اندازه های گوناگون در آن استفاده شده باشد.

ای یوسف خوش نام ما

خوش می روی بر نام ما

ای در شکسته جام ما

شکل ۱-۲: قلم های مختلف

## ۲-۳-۲ خواندن نویسه‌های فارسی و انگلیسی

در میان متون فارسی، عبارات انگلیسی زیاد دیده می‌شوند. الفبا هر دو زبان را در کنار هم می‌خواند و نویسه‌ها را با توجه به متفاوت بودن جهت نوشتار آن‌طور می‌چیند که متن به صورت یکپارچه خوانده شود. همچنین خواندن متون به زبان‌های عربی و دری نیز برای الفبا میسر است.

قوانینی که قرن‌ها  
پا برجا مانده است

Laws that have  
endured for centuries

شکل ۲-۲: متون فارسی و انگلیسی

## ۳-۳-۲ تحلیل ساختار سند

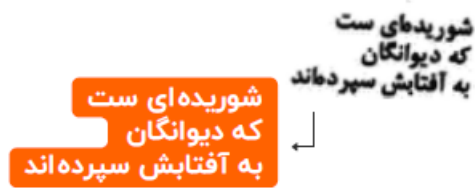
کلمه‌های فارسی در کنار هم می‌آیند و خطوط را شکل می‌دهند. خطهای پیاپی هم، پاراگراف را می‌سازند. این مفاهیم که در نوشته‌ها واضح‌اند، در تصویر سند از دست می‌روند و تبدیل به فاصله‌ها و طول‌ها می‌شوند. الفبا همین فاصله‌ها را تحلیل می‌کند و ساختار سند، از جمله خطوط و پاراگراف‌ها را پیدا می‌کند. همچنین محل عکس‌های سند در خروجی حاصل از تحلیل آن مشخص می‌شود.



شکل ۳-۲: ساختار سند

## ۴-۳-۲ اصلاح چرخش و نويز تصوير در سند

گاهی تصویر اسکن شده متن، با نویز همراه است یا ممکن است کاغذ آن کمی چرخیده باشد. این اشکالات پیش از تحلیل متن شند، برطرف می‌شوند.



شکل ۲-۴: اصلاح چرخش سند

## ۴-۲ اطلاعات تماس شرکت راهکار پردازش ژرف

آدرس : تهران، نارمک، خیابان فرجام، خیابان حسینعلی، پلاک ۹، واحد ۲۴  
شماره تلفن : ۰۲۱۹۱۰۰۳۳۰۳



## فصل سوم

### فعالیت‌ها و تجربیات کارآموزی



### ۱-۳ شرح موضوع کارآموزی

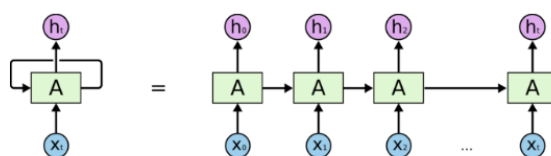
در نیمه‌ی اول این دوره‌ی کارآموزی تمرکز فعالیت‌ها روی یک پروژه آزمایشی با عنوان "نرمال‌سازی زبان محاوره با ابزار هضم" بود و پس از آن بخش اصلی دوره یعنی تولید داده‌های قطعه بندی <sup>۱</sup> برای سند هویتی کارت ملی آغاز شد که مورد نیاز آموزش مدل‌های جدید در پروژه الفبا بود. در ۲ بخش مجزا به این فعالیت‌ها می‌پردازیم.

#### ۱-۱-۳ پروژه نرمال‌سازی زبان محاوره

در این پروژه سعی می‌کنیم تا متون محاوره زبان فارسی را نرمال‌سازی کنیم و آنها را به صورت رسمی تبدیل کنیم. برای این منظور ابتدا ابزارهای قرار داده شده برای تحلیل زبان محاوره در هضم را بررسی می‌کنیم و محدودیت‌های آنها را مشخص می‌کنیم. در گام بعد مقاله‌ی مرتبط با همین موضوع <sup>[۱]</sup> را بررسی می‌کنیم و نتایج آنرا تا جای ممکن بازتولید می‌کنیم. در نهایت اقدامات قابل انجام در پروژه هضم را مشخص می‌کنیم و گام به گام آنها را اجرا می‌کنیم. در ابتدا به صورت خلاصه مطالبی که درباره موضوع این پروژه آموخته ام بیان می‌کنم. گزارش‌های مرتبط با این پروژه در بخش پیوست آمده است.

#### ۱-۱-۱-۳ شبکه عصبی بازگشتی

شبکه‌های عصبی بازگشتی (RNN) یک لایه بازخورد دارند که در آن خروجی شبکه به همراه ورودی بعدی، به شبکه بازگردانده می‌شود. RNN می‌تواند به علت داشتن حافظه داخلی، ورودی قبلی خود را به خاطر بسپارد و از این حافظه برای پردازش دنباله‌ای از ورودی‌ها استفاده کند. به بیان ساده، شبکه‌های عصبی بازگشتی شامل یک حلقه بازگشتی هستند که موجب می‌شود اطلاعاتی را که از لحظات قبلی بدست آورده ایم از بین نروند و در شبکه باقی بمانند.



شکل ۱-۳: شبکه عصبی بازگشتی

<sup>۱</sup>Generate segmentation data

## ۳-۱-۱-۲ مدل زبانی

مدل کردن زبان تسکی<sup>۲</sup> است که در آن می‌آییم و کلمه بعدی را پیش‌بینی می‌کنیم؛ یعنی دنبال‌ای از کلمات را داریم و می‌خواهیم توزیع احتمال کلمه بعدی در جمله را بدست آوریم.

$$P(x^{(t+1)} | x^{(t)}, \dots, x^{(1)}) \quad (۱-۳)$$

در ادامه درباره ترجمه ماشین صحبت می‌کنیم که یک تسک جدید در پردازش زبان طبیعی است و در آن از معماری عصبی جدید sequence-to-sequence استفاده می‌شود و در آخر هم اشاره کوتاهی به تکنیک عصبی attention می‌کنیم که seq2seq<sup>۳</sup> را بهبود می‌دهد.

## ۳-۱-۱-۳ تاریخچه ترجمه ماشین

به کار ترجمه کردن جمله x از زبان مبدا به جمله y در زبان مقصد، تسک machine translation می‌گویند. آغاز ترجمه ماشین به عنوان یک تسک در هوش مصنوعی به سال ۱۹۵۰ برمی‌گردد. در زمان جنگ سرد، غرب علاقه داشت تا زبان روسی را به انگلیسی ترجمه کند و از اطلاعات استفاده کند. اما در دهه بعد ترجمه ماشینی آماری<sup>۴</sup> مطرح شد که ایده اصلی آن یادگیری یک مدل احتمالی از داده‌ها بود. در آن مفاهیم Translation model و Language model بیان شدند. مدل ترجمه نشان می‌دهد که چگونه کلمات را باید با استفاده از آموزش از طریق داده‌های موازی موسوم به parallel data ترجمه کنیم. همانطور که پیشتر بیان کردیم، مدل زبان هم سیستمی است که می‌تواند کلمه بعدی را پیش‌بینی کند [۲].

## ۳-۱-۱-۴ ترجمه ماشینی عصبی

ایده اولیه آن این بود که از طریق یک شبکه عصبی واحد فرآیند ترجمه صورت بگیرد. معماری شبکه عصبی که استفاده می‌شود، seq2seq<sup>۵</sup> است که از RNN<sup>۲</sup> تشکیل شده است. نام این معماری seq2seq است؛ به این دلیل که یک دنباله<sup>۶</sup> را به دنباله‌ای دیگر متصل می‌کند. همچنین به RNN<sup>۲</sup> نیاز داریم تا جمله را بررسی کنیم.

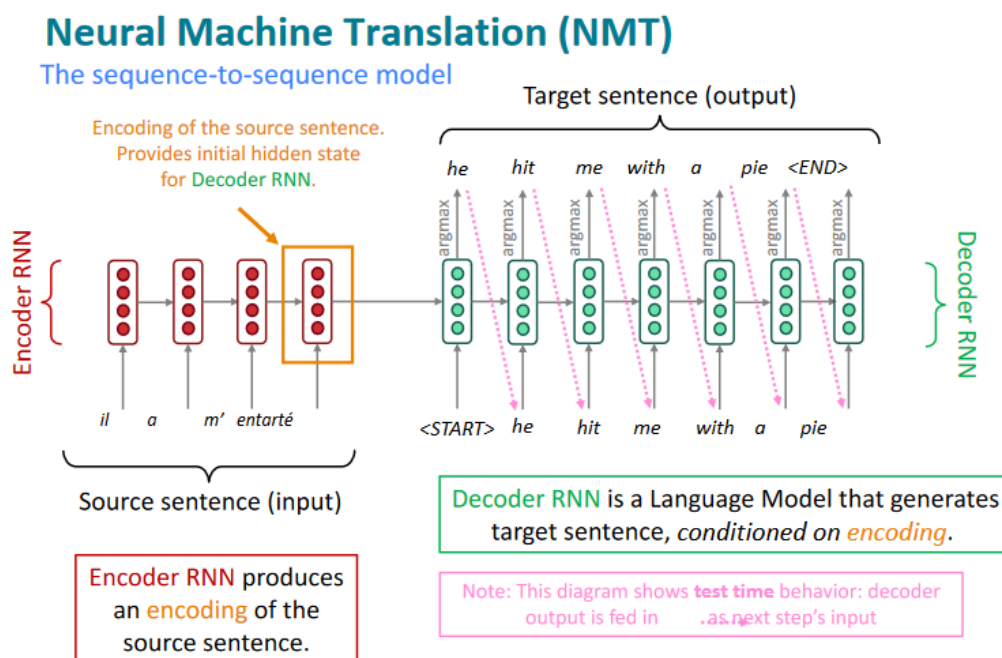
<sup>۲</sup>Task<sup>۳</sup>sequence-to-sequence<sup>۴</sup>Statistical Machine translation (SMT)<sup>۵</sup>sequence

## ۳-۱-۵ مدل seq2seq

این شبکه‌ها از ۲ بخش کدگذار (Encoder) و کدگشا (Decoder) تشکیل شدند. کدگذاری که می‌توانیم استفاده کنیم هر کدام از انواع مختلف RNN<sup>۶</sup> می‌تواند باشد.

بخش کدگذار، یک سری می‌گیرد و یک بردار به عنوان خروجی برمی‌گرداند و بخش کدگشا برعکس. کدگذار جمله از زبان مبدأ را به یک بردار تبدیل می‌کند و قسمت کدگشا این بردار را می‌گیرد و تبدیل به جمله در زبان مقصد می‌کند.

برای آموزش این شبکه، جمله زبان مبدأ به Encoder داده می‌شود و جمله مقصد به همراه یک شیفت زمانی به Decoder داده می‌شود؛ یعنی در هر مرحله کلمه‌ای را به Decoder می‌دهیم که باید در مرحله قبل خروجی می‌داد و به عنوان اولین ورودی (قبل از اولین کلمه جمله) به آن یک توکن  $\langle \text{start} \rangle$  می‌دهیم و جمله مبدأ را هم برعکس می‌کنیم که اولین کلمه‌ی جمله آخرین ورودی به Encoder است. به صورت معمول این برعکس کردن انجام نمی‌شود. و پس از مدتی عملکرد ضعیف این شبکه‌ها، این ایده به ذهن محققین رسید که جمله ورودی را برعکس به Encoder بدهند. این باعث می‌شود که کلمات ابتدایی جمله‌ی مبدأ فاصله کم‌تری با کلمات ابتدایی جمله‌ی مقصد داشته باشد. حال این ریسک<sup>۸</sup> «فاصله کم‌تر» می‌تواند باعث تولید شدن بهتر کلمات ابتدایی جمله مقصد بشود و این امید هست که در ادامه هم جمله مقصد بهتری تولید شود.



شکل ۳-۲: مدل sequence-to-sequence

<sup>۶</sup>uni-directional, bi-directional, multi-layer, LSTM, ...

<sup>۷</sup>Token

<sup>۸</sup>Risk

## ۳-۱-۶ مکانیزم توجه (attention)

اما مشکلی که داریم مسیر طولانی است که از اولین Encoder تا آخرین Decoder باید بپیمائیم. کلمه باید در تمام این مسیر باید کامل و سالم منتقل شود بدون اینکه احتمالاً استفاده‌ای از آن بشود و همینطور باید گرادیان تمام این مسیر را هم برگردد. خب نمیشود این مسیر را کوتاه‌تر کنیم؟ می‌شود. اینجاست که «مکانیزم توجه» ورود می‌کند. مکانیزم توجه به این صورت است که در آن ساختار کدگشا می‌تواند در هر قدم فقط روی کلمات مناسب تمرکز و «توجه» کند. این ایده باعث شد مسیر بین یک کلمه در زبان مرجع و ترجمه‌اش خیلی کوتاه‌تر بشود و مشکل حافظه‌ی کوتاه مدت شبکه‌های بازگشتی خیلی کم‌تر مشکل ساز بشود.

## ۳-۱-۲ پروژه تولید داده‌های قطعه‌بندی

بعد از پروژه اولیه‌ای که به من سپرده شد، از من درخواست شد تا پروژه دیگری را پیش بگیرم؛ چرا که به میزان کافی برای پروژه اولیه که بیشتر حکم آشنایی را داشت وقت گذاشته بودم. پروژه جدید مربوط به ابزار الفبا می‌شد. ویژگی‌های این ابزار در فصل ۲ بیان شد. وظیفه‌ای که به من سپرده شد این بود که کد مربوط به این محصول را بخوانم و مرور <sup>۹</sup>کنم و در ادامه آنرا گسترش دهم. من می‌بایست کدی بنویسم که یک سند هویتی مثل کارت ملی تولید می‌کرد. در ابتدا برای مدل segmentation به دیتاستی از کارتهای ملی نیاز بود که باید تولید می‌شدند و در ادامه می‌بایست augmentation هایی روی تصاویر اعمال می‌کردم از قبیل:

- اعمال طیف سایه برای تصاویر
- تغییر جهت کارت ملی
- اعمال نویز تار شدن تصاویر
- اعمال نویز برجسته شدن کارت ملی
- و غیره

گزارش آخر در بخش پیوست مربوط به این پروژه است که جزئیات انجام این پروژه را در آنجا نوشته‌ام. با انجام این تسک، کار من به عنوان کارآموز در شرکت روشن به اتمام رسید.

<sup>9</sup>Review

## فصل چهارم

### جمع‌بندی و نتیجه‌گیری

## ۴-۱ نتیجه‌گیری

هدف این کارآموزی آشنایی با تکنولوژی‌های موجود در حوزه هوش مصنوعی بود و من به عنوان کارآموز در حوزه‌های ترجمه ماشینی Neural Machine Translation و تولید داده‌های قطعه‌بندی Generate Segmentation Data فعالیت کردم. این کار با تمام سختی‌های بسیاری که داشت، اما با وجود تیم حرفه‌ای و همکاران باتجربه‌ی حاضر در شرکت به خوبی انجام شد. سرعت یادگیری مفاهیم جدید در کنار اعضای باتجربه شرکت به شدت بیشتر از مطالعه فردی بود و مزیت اصلی آن امکان پرسش و پاسخ با همکاران بود. به عنوان کارآموز هوش مصنوعی، از حضور در این جایگاه رضایت حاصل می‌شود؛ زیرا علاوه بر چالش‌های موجود در توسعه، به علت رسیدن نتیجه‌ی زحمات به کاربران در آینده، احساس بهتری نیز به افراد می‌دهد.

## ۴-۲ کارهای آینده

پروژه الفبا توسعه خود را ادامه خواهد داد و با تولید داده‌های جدید مدلها آماده آموزش می‌شوند. به احتمال زیاد در آینده زمان بیشتری برای اضافه کردن augmentatin ها و تولید سندهای هویتی جدید گذاشته خواهد شد.

## منابع و مراجع

- [1] Rasooli, Mohammad Sadegh, Bakhtyari, Farzane, Shafiei, Fatemeh, Ravanbakhsh, Mahsa, and Callison-Burch, Chris. Automatic standardization of colloquial persian. arXiv preprint arXiv:2012.05879, 2020.
- [2] Manning, Christopher. Natural language processing with deep learning cs224n/ling284, April 2019.

# پیوست

## گزارش‌های کارآموزی

در این قسمت گزارش‌های دوره‌ای که در طول مدت کارآموزی برای دو پروژه نوشته شده است، می‌آید. تمام طول مدت کارآموزی شامل ۳ بخش زیر می‌باشد:

- آموزش و تکمیل فراگیری مباحثی از یادگیری ماشین و پردازش زبان طبیعی
- آشنایی با کد پروژه الفبا بخش داده‌های قطعه‌بندی موسوم به alefba-segmentation
- پیاده‌سازی وظایف محوله

این گزارش‌های دوره‌ای بنا به درخواست مسئول کارآموزی شرکت روشن نگاشته شده است و به این دلیل که مستند و دقیق است، همان گزارش‌ها در اینجا آورده شده است. لازم به ذکر است که این گزارش‌ها شامل ۴ بخش زیر می‌باشند که قالب کلی برای نگارش گزارش‌ها را مشخص می‌کند.

۱. توضیح مختصر درباره‌ی آموزش‌های روزانه

۲. اقدامات انجام شده در آن روز

۳. برنامه اقدامات فردا

۴. مشکلات پیش‌آمده



## گزارش شماره ۱

تاریخ: ۰۰/۰۵/۰۲

## توضیح مختصر درباره‌ی آموزش‌های امروز:

از این هفته با تمام شدن ترم دانشگاه به صورت جدی تر مطالعات را شروع کردم و از سر گرفتم. در ابتدا مطالعه ۲ منبع را شروع کرده ام و در حال مرور مطالب مرتبط با مقاله هستم. ۲ منبع به شرح زیر هستند:

- درس پردازش زبان طبیعی دانشگاه استنفورد<sup>۱</sup>

- درس یادگیری ماشین دکتر رهبان دانشگاه شریف<sup>۲</sup>

درس پردازش زبان طبیعی را یکی از اساتید با تجربه دانشگاه استنفورد تدریس می‌کند. عنوان این درس NLP with Deep learning هست و در سال ۲۰۱۹ از کلاس فیلمبرداری شده است. درس یادگیری ماشین را هم یکی از اساتید برجسته دانشگاه صنعتی شریف تدریس کرده‌اند و در اختیار عموم قرار داده‌اند. این کلاس در سال ۱۳۹۹ از کلاس مجازی فیلمبرداری شده است.

## اقدامات انجام شده امروز:

- مشاهده فیلم‌های آموزشی
- حل برخی از مسائل مطرح شده در فیلمها
- به سراغ مقاله رفتم و کمی انرا مطالعه کردم و نکات مهم آنرا بررسی کردم. و همچنین کد مربوط به مقاله را از سایت [paperswithcode](https://paperswithcode.com) پیدا کردم و سعی کردم آنرا اجرا کنم اما به چند خطا برخوردم.

## برنامه اقدامات فردا:

فردا قصد دارم تا به سراغ مستندات ابزار هضم (محصول شرکت راهکار پردازش ژرف) بروم؛ زیرا که قرار است در ادامه از این ابزار استفاده کنم. مستندات هضم از این آدرس<sup>۳</sup> قابل دسترسی هست.

## مشکلات پیش آمده:

- مشکل خاصی نداشتم اما در ابتدا مفهوم Skip-grams در بحث Word2vec را به خوبی درک نکرده بودم که با کمی جستجوی بیشتر مشکلم برطرف شد.

<sup>۱</sup><https://bit.ly/3ufToBb>

<sup>۲</sup><https://www.aparat.com/v/bWr3A?playlist=954603>

<sup>۳</sup><https://www.sobhe.ir/hazm/docs/>

- کمی در مورد موضوع تسک<sup>۴</sup> اصلی ام فکر کردم و ابهام‌هایی برایم ایجاد شد که باید در جلسه آنلاین بعدی از مسئول کارآموزی شرکت بپرسم.

---

<sup>4</sup>Task

## گزارش شماره ۲

تاریخ : ۰۰/۰۵/۰۳

## توضیح مختصر درباره‌ی آموزش‌های امروز :

- امروز دوره پردازش زبان طبیعی دانشگاه استنفورد را ادامه دادم و فیلمهای مرتبط تر با موضوع اصلی تسک را مشاهده کردم.
- یک دوره رایگان به زبان فارسی با تدریس استاد مطرح در زمینه پردازش زبان از سایت [bigdataworld](https://bigdataworld.com) با عنوان "آموزش پردازش زبان طبیعی با یادگیری عمیق" پیدا کردم که مطالب را به صورت کاربردی بیان می‌کرد. من هم همگام با ایشان پیاده‌سازی‌ها را انجام دادم و برخی مطالب را آموختم و برخی مطالب برایم مرور شد.

## اقدامات انجام شده امروز :

- مشاهده فیلمهای دانشگاه استنفورد و سایت جدید بیگ دیتا
- در جلسه‌ای که امروز داشتیم، یکی از اعضای تیم در مورد زبان markdown محیطی را به نام Typora معرفی کردند. ایشان گفتند که نوشتن متن مارک‌داون با این برنامه خیلی راحت هست.

## برنامه اقدامات فردا :

- در جلسه ای که امروز با مسئول کارآموزی داشتم، سوالاتم از تسک اصلی را پرسیدم. به علاوه ایشان انتظار داشتند تا گزارش‌ها را در محیط Gitlab در issue ای که برای من مشخص شده است بنویسم. اما از آنجا که با markdown زیاد آشنا نیستم، باید فردا کمی در مورد نحوه کار کردن با این زبان تحقیق کنم.

## مشکلات پیش آمده :

امروز مشکلی نداشتم.

## گزارش شماره ۳

تاریخ : ۰۰/۰۵/۰۴ تا ۰۰/۰۵/۲۵

توضیح مختصر درباره‌ی آموزش‌های امروز :

متأسفانه گزارش‌های روزهای ۵ مرداد الی ۲۴ مرداد را به دلیل بیماری نتوانستم به صورت تفکیک شده بنویسم و در جلسات هم به صورت نامنظم شرکت کردم. اما به صورت تجمیعی در این مدت به کارهای زیر پرداختم:

- تکمیل مشاهده دوره‌های پردازش زبان طبیعی دانشگاه استنفورد و بیگ‌دیتا
- از سایت kaggle دوره‌های کوتاهی که داشت را مشاهده کردم و تمرینش را حل کردم. این دوره‌ها در مورد مفاهیم pandas, natural language processing, machine learning بود.
- دوره nlp specialization در سایت coursera را هم دیدم و مباحث sequence و attention را مطالعه کردم.

اقدامات انجام شده امروز :

- مشاهده فیلم‌های آموزشی
- مطالعه مقاله ضمیمه شده به تسک [۱] و جمع آوری dataset های مختلفی که در مقاله استفاده شده است. از جمله آنها دیتاست های مقابل هستند:

TEP \*

Mizan \*

wikipedia dump \*

Shekasteh \*

- مطالعه مقاله معروف و مهم attention is all you need و همچنین مطالعه ۲ مطلب مرتبط با آن

\* لینک ۱

\* لینک ۲

برنامه اقدامات فردا :

- بررسی مشکلات ابزار هضم

- نوشتن قطعه کدهایی برای خواندن دیتاستها. بعضی از دیتاست ها مثل Shekasteه فرمت tsv دارد و نمیدانم برای خواندنش باید چه طور اقدام کنم.
- اجراکردن مجدد کد پیاده سازی شده ی مقاله و رفع ایرادات آن

#### مشکلات پیش آمده :

- مطمئن نیستم که آیا باید کد پیادسازی مقاله را حتما اجرا کنم یا خیر. چرا که به نظر می رسد از ابتدا خودم باید کد جدیدی برای این تسک بنویسم.

## گزارش شماره ۴

تاریخ : ۰۵/۲۶ تا ۰۶/۰۲/۰۰

توضیح مختصر درباره‌ی آموزش‌های امروز :  
امروز آموزشی ندیدم و صرفاً به دنبال پیاده‌سازی ها بودم.

اقدامات انجام شده امروز :

در این مدت یک هفته، تسک ساده‌تر از پروژه اصلی برای خودم تعریف کردم و در خلال آن به مشکلاتی برخوردیم که جلوتر بیان می‌شود. در ابتدا باید بیان کنم که تسک اصلی من در مورد ترجمه هست که در حوزه پردازش زبان با Neural Machine Translation شناخته می‌شود. تسک جدید و کوچکی که برای خودم تعریف کردم هم درباره همین ترجمه بود که می‌آمد و یک جمله در زبان پرتغالی را به زبان انگلیسی ترجمه می‌کرد. مجموعه دادگان<sup>۵</sup> برای این تسک را از سایت [manythings.org](http://manythings.org) دریافت کردم و در Google Drive ام ذخیره کردم تا راحت تر بتوانم با آن کار کنم. حجم این دیتاست در حدود ۲۷ مگابایت است و فرمت آن txt می‌باشد. نمونه‌ای از آنرا در شکل<sup>۵</sup> می‌بینید.

	source	target
107683	I wanted Tom to stay with Mary.	Eu queria que o Tom ficasse com a Mary.
107684	I wanted time instead of money.	Eu queria tempo em vez de dinheiro.
107685	I wanted time instead of money.	Queria tempo em vez de dinheiro.
107686	I wanted to ask you a question.	Eu queria te fazer uma pergunta.
107687	I wanted to be a piano teacher.	Eu queria ser professor de piano.
107688	I wanted to do the right thing.	Eu queria fazer a coisa certa.
107689	I wanted to get Tom to help me.	Eu queria receber o Tom para me ajudar.
107690	I wanted to go, but I got sick.	Eu queria ir, mas fiquei doente.
107691	I wanted to know what happened.	Eu queria saber o que aconteceu.
107692	I wanted to know where Tom was.	Eu queria saber onde Tom estava.
107693	I wanted to play golf with Tom.	Eu queria jogar golfe com o Tom.
107694	I wanted to say goodbye to Tom.	Eu queria me despedir de Tom.
107695	I wanted to say that I'm sorry.	Eu queria dizer que sinto muito.
107696	I wanted to speak to Tom alone.	Eu queria falar com Tom sozinho.
107697	I wanted to stay home with Tom.	Eu queria ficar em casa com Tom.
107698	I warned you this would happen.	Eu te avisei que isso poderia acontecer.
107699	I warned you this would happen.	Te avisei que isso poderia acontecer.

شکل ۱: مجموعه دادگان تسک اولیه

این دیتاست دارای ۱۷۷۷۴۱ نمونه است و دارای ۲ ستون می‌باشد که یک ستون مبدا و دیگری مقصد

<sup>5</sup>dataset

هست. در حقیقت به این دیتاست parallel corpus می‌گویند. در این پروژه از مکانیزم attention که هفته گذشته با آن آشنا شدیم استفاده شد و encoder و decoder هایی تعریف شدند و همچنین به دلیل راحتی به جای LSTM از GRU استفاده شد.

- پیاده‌سازی تسک کوچک اولیه که برای خودم تعریف کرده بودم
- مقایسه ابزار هضم (از محصولات شرکت راهکار پردازش ژرف) با ابزار فارسیار. قرار بود که در این تسک الگوی خاصی درباره‌ی تبدیل زبان محاوره به رسمی در خطاهای هضم پیدا کنم؛ اما فقط توانستم چند مورد را پیدا کنم که در شکل ۱ مشخص است.


اصلی	فارسیار	هضم
اگه اون گزینه رو کلیک کنی، به پنجره باز میشه که میتونی رمز عبورتون رو اونجا تغییر بدی داشتیم ی رفتم برم، دیدم گرفت نشست، گفتم بذار پرسم ببینم میاد نمیا دادم میگه نمیخوام پیام بذار برم بگیرم بخوابم نمیتونم بشینم کتابای خودتونه نمیدونم چی بگم که دیگه اونجا نره ساعت جن میتونن بیاین؟	اگر آن گزینه را کلیک کنید، یک پنجره باز می‌شود که می‌توانید رمز عبورتان را آنجا تغییر بدهید داشتیم می‌رفتم بروم، دیدم گرفت نشست، گفتم بگذار پرسم ببینم می‌آید نمی‌آید دیدم می‌گوید نمی‌خواهم پیام بگذار بروم بگیرم بخوابم نمی‌توانم بنشینم کتاب‌های خودتان است نمی‌دانم چه بگویم که دیگر آنجا نرود ساعت چند می‌توانید بیایید؟	اگر آن گزینه رو کلیک کنی ، یک پنجره باز می‌شود که می‌توانی رمز عبورتان رو اونجا تغییر بدی داشتیم می‌رفتم برم ، دیدم گرفت نشست ، گفتم بذار پرسم ببینم میاد نمیا دادم می‌گویند نمی‌خواهم پیام بذار برم بگیرم بخوابم نمی‌توانم بشینم . کتابای خودتونه نمی‌دانم چه بگویم که دیگر اونجا نرود ساعت جن می‌توانی بیایند ؟


جدول ۱: مقایسه ابزار هضم و ابزار فارسیار با متن اصلی

همانطور که مشاهده می‌شود، خروجی ابزار فارسیار با ورودی تصادفی بهتر از هضم عمل کرد. در ادامه الگوهایی از خطاهای هضم پیدا شدند و درصد رفع آنها برآمدم. همچنین در ادامه این ۲ ابزار را به صورت کوتاه معرفی کردم.

◇ هضم کتابخانه‌ای است که در حوزه پردازش متن فارسی استفاده می‌شود. از ویژگی‌های مهم آن می‌توانیم به این موارد اشاره کنیم:


✱ تمیز و مرتب کردن متن

✱ قطعه‌بندی متن 

✱ نشانه‌گذاری 

✱ سازگاری با بسته NLTK

✱ ریشه‌یابی واژه‌ها

◇ فارسیار هم ابزار دیگری است که در متن‌کاوی فارسی مورد استفاده قرار می‌گیرد و از طریق سایت‌شان در دسترس هست. 

<sup>6</sup>Sentence segmentation

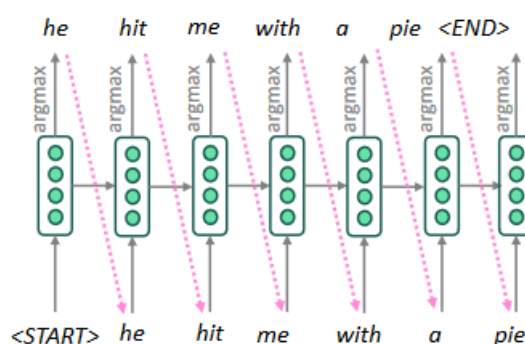
<sup>7</sup>Tokenization

<sup>8</sup><https://app.text-mining.ir/>

- همچنین کد مقاله را اجرا کردم و هنوز نتوانستم مشکلات آنرا رفع کنم. نویسندگان مقاله کدی را که پیاده‌سازی کرده‌اند ابزار PBreak نامیده‌اند که مخفف Persian Break words می‌باشد و به معنای کلمات شکسته در زبان فارسی می‌باشد.

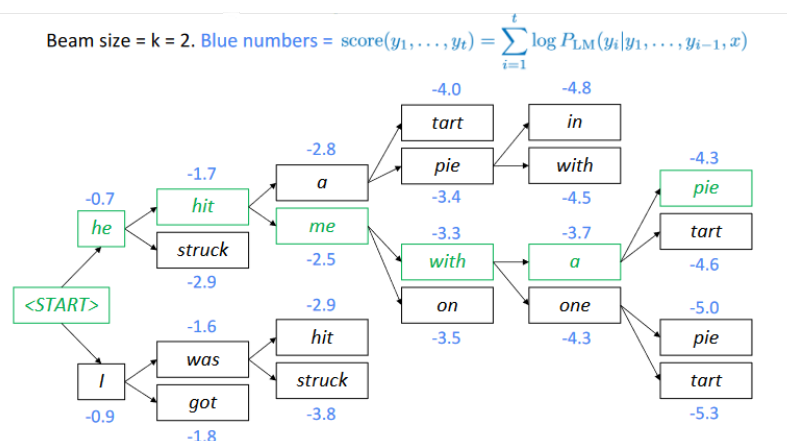
مقایسه Greedy decoding و Beam-search decoding

- می‌توانیم با محاسبه  $\text{argmax}$  از word vector ها بیاییم و بهترین کلمه را انتخاب کنیم و به مرحله بعدی برویم و کلمه جدید را به state بعدی decoder وارد کنیم. در این صورت دیگر راهی برای به عقب برگشتن نداریم و به همین دلیل به این روش greedy decoding و یا کدگشایی حریصانه می‌گویند. در شکل ۲ مثالی از این کدگشایی را می‌بینیم.



شکل ۲: کدگشایی حریصانه

- ایده دیگر beam-search است که در هر مرحله از decoding ،  $k$  تا از محتمل ترین ترجمه های جزئی را پیگیری می‌کند که به آن hypotheses می‌گویند. مقدار  $k$  ، beam-size می‌باشد. در شکل ۳ مثالی از این کدگشایی را می‌بینیم.



شکل ۳: کدگشایی جستجوی پرتو محلی

برنامه اقدامات فردا :



- در مقاله جایی آورده بود که از مقاله‌ای تحت عنوان "صورت‌های شکسته و فرهنگ‌های فارسی" از آقای امید طبیب‌زاده استفاده شده است و باید آنرا پیدا کنم.

### مشکلات پیش آمده :

- مشکلی برایم پیش آمد این بود که epoch ها بشدت کند اجرا می‌شدند به صورتی که برای هر epoch نزدیک به ۴۰ دقیقه منتظر می‌ماندم. لازم به ذکر است که در محیط google colab کدها را اجرا کردم و از GPU هم استفاده کردم اما فرق چندانی نداشت.
- مشکلی دیگری که داشتم در مورد اجرای کد PBREAK بود و کتابخانه wapiti نصب نمی‌شد.
- در مورد متن مقاله هم به ابهاماتی برخورد کردم. در مقاله از چندین دیتاست نام برده شده بود و کمی گیج‌کننده بود که از کدام به عنوان دیتاست اصلی باید استفاده کرد. این طور که تا بدینجا متوجه شدم این است که از چندین دیتاست در آن پروژه استفاده شده است. نام این دیتاستها در گزارشهای قبلی آمده‌اند.
- همچنین در مقاله آمده است که از الگوریتمی استفاده شده است؛ اما من چنین الگوریتمی که توصیف شد را در بین کدهای مقاله ندیدم.
- در جای دیگری از مقاله آمده است که از greedy decoding استفاده شده است در حالیکه کمی جلوتر بیان شده که از beam search استفاده شده است. و من متوجه نشدم که باید از کدام استفاده کنم.


## گزارش شماره ۵

تاریخ : ۰۰/۰۶/۰۳ تا ۰۰/۰۶/۰۵

## توضیح مختصر درباره‌ی آموزش‌های امروز :

در این چند وقت مطالعه‌ای نداشتم و صرفاً در حال پیاده‌سازی و رفع نواقص بودم. البته مقالاتی در حین کار مطالعه کردم که بیشتر در حوزه مفاهیم epoch و batch-size و iteration بود که تفاوت‌ها و ارتباطشان را بیان می‌کرد.

## اقدامات انجام شده امروز :

همانطور که هفته گذشته بیان کردم، در مقاله جایی آورده بود که از مقاله‌ای تحت عنوان "صورت‌های شکسته و فرهنگ‌های فارسی" از آقای امید طبیب‌زاده استفاده شده است که به سختی توانستم مقاله‌شان را پیدا کنم. 

## برنامه اقدامات فردا :

باید زیرمجموعه‌ای از دیتاست را انتخاب کنم چرا که همچنان برنامه طول می‌کشد. تصمیم گرفتم که در ادامه با همان دیتاست اصلی مقاله یعنی Shekasteh جلو بروم.

## مشکلات پیش آمده :

همچنان مشکل طول کشیدن بیش از حد ادامه دارد. طی جلسه‌ای که با مسئول کارآموزی شرکت داشتم متوجه شدم که باید زیرمجموعه‌ای از دیتاست را تست کنم تا دریابم که آیا مشکل از مدلی است که طراحی کرده‌ام یا خیر.

---

<sup>9</sup><https://bit.ly/2XMWSPE>

## گزارش شماره ۶

تاریخ : ۰۰/۰۶/۰۶ تا ۰۰/۰۶/۰۷

## توضیح مختصر درباره‌ی آموزش‌های امروز :

• ابتدایه سراغ آن رفتیم تا مدلم را بهبود دهیم. و در ادامه به سراغ معنای دقیق BLEU-score رفتیم و مطالعه داشتیم. و همچنین در تفاوت LSTM و GRU ابهام داشتیم که برطرف شد. در ادامه انواع attention را بررسی کردم و مکانیزمهای توجه باهدانا<sup>۱۰</sup> و لوآنگ<sup>۱۱</sup> را به طور ویژه بررسی کردم چراکه در مقاله از آنها نام برده شده بود.

• برای ارزیابی سامانه مترجم<sup>۱۲</sup> از معیار BLEU-score استفاده می‌کنیم که مخفف Bilingual Evaluation Understudy می‌باشد. نحوه بدست آوردن این مقدار را با یک مثال نشان می‌دهیم. ابتدا مفهوم BLEU-score را در bigrams<sup>۱۳</sup> بیان می‌کنیم. فرضاً جمله‌ای به زبن فرانسوی را به مدل می‌دهیم و در خروجی ترجمه‌ای را به ما برمی‌گرداند. همچنین از ۲ انسان هم می‌خواهیم تا این جمله فرانسوی را ترجمه کنند. حالا ۲ ترجمه انسانی و یک ترجمه ماشینی داریم.

French sentence: Le chat est sur le tapis.

Reference 1: The cat is ont the mat.

Reference 2: There is a cat on the mat.

Machine translation output: The cat the cat on the mat.

حال می‌آییم و جفت کلمه‌های کنار یکدیگر که در جمله ترجمه شده توسط ماشین وجود دارند را جدا می‌کنیم و تعدادشان را محاسبه می‌کنیم. همه این مقادیر را با یکدیگر جمع می‌کنیم و به عنوان مخرج کسری که جلوتر بیان می‌کنیم در نظر می‌گیریم.

از طرفی این جفت کلمه‌ها را با جفت کلمه‌های موجود در ترجمه‌های انسانی مقایسه می‌کنیم و بیشترین تعداد تکرار را انتخاب می‌کنیم. همه این مقادیر را با یکدیگر جمع می‌کنیم و به عنوان صورت کسر در نظر می‌گیریم.

برای مثال بالا عدد  $\frac{4}{5}$  را بدست خواهیم آورد. همین کار را برای unigram و به طور کلی n-gram

<sup>10</sup>Bahdanau

<sup>11</sup>Luong

<sup>12</sup>Machine Translation System

<sup>۱۳</sup>جفت کلمه‌ها

هم انجام می‌دهیم. و در نتیجه معیار دقت اصلاح شده <sup>۱۴</sup> را تعریف می‌کنیم:

$$p_n = \frac{\sum_{C \in \{\text{Candidates}\}} \sum_{n\text{-gram} \in C} \text{Count}_{\text{clip}}(n\text{-gram})}{\sum_{C' \in \{\text{Candidates}\}} \sum_{n\text{-gram}' \in C'} \text{Count}(n\text{-gram}')} \quad (۱)$$

$$BLEU\text{score} = BP * \exp\left(\frac{1}{n} \sum_{i=1}^n p_n\right) \quad (۲)$$

مقدار BP برای تعدیل این معیار ضرب می‌شود و به این صورت محاسبه می‌شود.

$$\text{Brevity Penalty} = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases} \quad (۳)$$

r به معنای تعداد کلمات در ترجمه مرجع است و c به معنای تعداد کلمات در ترجمه ماشین است.

اقدامات انجام شده امروز :

- در این چند روز بیشتر مطالعه داشتیم.
- در مورد پیاده‌سازی دیدم که وقتی تعداد نمونه‌ها از حدود ۱۷۷۰۰۰ به ۲۰۰۰ کاهش داده‌شد، به جای اینکه هر epoch ۵۰ دقیقه طول بکشد، هر iteration حدودا ۵۰ دقیقه طول میکشد و روند آموزش مدل سری‌تر شد. اما همچنان دقت خوبی ندارم و ترجمه هم به خوبی صورت نمی‌گیرد و مقدار BLEU-score در حدود ۷.۰ بود.

برنامه اقدامات فردا :

- طبق صحبتی که با مسئول کارآموزی داشتیم، قرار شد تا دیتاست را بر اساس طول جملات دسته‌بندی کنم و آن نمونه‌هایی که طول جملات کوتاه‌تری دارند را برای بحث آموزش مدل انتخاب کنم.

مشکلات پیش آمده :

مشکلی در مفاهیم نداشتیم.

<sup>14</sup>modified precision score



## گزارش شماره ۸

تاریخ : ۰۰/۰۶/۱۰ تا ۰۰/۰۶/۱۲

توضیح مختصر درباره‌ی آموزش‌های امروز :

در این مدت آموزش خاصی ندیدم اما لازم می‌دانم تا مفاهیمی مثل epoch ، batch-size و iterations را با همدیگر مقایسه کنم و در این گزارش بیاورم.

اقدامات انجام شده امروز :

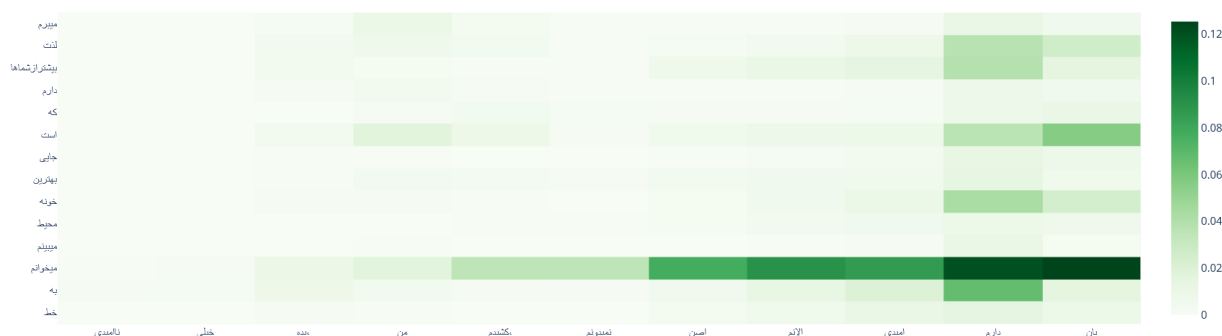
در این چند روز، خیلی پیشرفت خوبی نداشتم ولی به هر حال کدهای مربوط به Encoder و Decoder را مرتباً تغییر می‌دادم و همچنین مقادیر ثابت مثل batch-size و embedding-dimension را تغییر دادم اما پیشرفت خوبی دریافت نکردم. اما ترجمه‌ها تا حدی بهتر و دقیق‌تر شدند. جزئیات مربوط به تعداد epoch ها و batch-size به همراه خروجی مدل را در زیر مشاهده می‌کنیم. در ضمن دیتاست مورد بحث همان دیتاستی که در مقاله به کار رفته است، می‌باشد. در شکل ۵ خروجی مدل را در این مرحله برای یک متن تصادفی آورده‌ام.

Training examples (samples) = ۴۷۷

Number of epochs = ۷۰

Batch-size = ۶۴

Iterations = ۵۲



شکل ۵: نمایش خروجی مدل بهبودیافته

متنی ورودی به مدل : ناامیدی خیلی بده، من کشیدم، نمیتونم اصن الانم امیدو دارم یان.  
متن خروجی از مدل : خط به میخوانم میبینم محیط خونه بهترین جایی است که دارم بیشتر از شماها لذت میبرم.

برنامه اقدامات فردا :

هنوز بقیه اقداماتی که در مقاله بیان شده را پیاده‌سازی نکرده‌ام و در گام‌های ابتدایی هستم.

مشکلات پیش آمده :

تنها مشکل مربوط به مدل هست که همچنان نمیدانم باید چه اقدامی برای آن انجام دهم.

## گزارش شماره ۹

تاریخ : ۰۰/۰۶/۱۲ تا ۰۰/۰۶/۲۴

توضیح مختصر درباره‌ی آموزش‌های امروز :

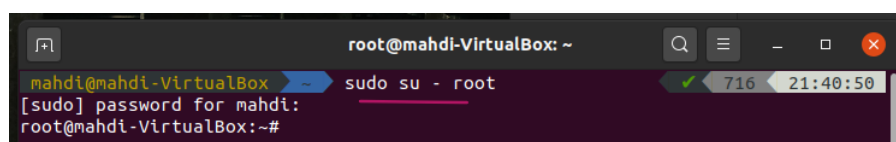
اقدامات انجام شده امروز :

- در گام ابتدایی به سراغ مطالعه کدها رفتیم و با ۲ نفر از اعضای تیم الفبا جلسه گذاشتیم.
- در ادامه سعی کردم تا ماشین مجازی لینوکس نصب کنم؛ چراکه تمام اعضای تیم با سیستم عامل لینوکس برنامه‌نویسی می‌کردند و کدها برای اجرا شدن با کتابخانه‌های مختص لینوکس سازگارتر بودند و بهتر و آسانتر اجرا می‌شدند.
- در ادامه سعی کردم تا تنظیماتی را انجام دهم تا به سرور شرکت متصل شوم. البته این فرآیند کمی برایم طول کشید. خلاصه این فرآیند را در ادامه خواهم نوشت. برای اینکه به سرور دسترسی داشته باشیم باید به یک VPN وصل بشویم ( که همین باعث می‌شود تا اینترنت کار نکند ) و بعد از آن باید به اصطلاح SSH بزنیم.

\* ساخت SSH key - باید با ssh-keygen جلو برویم و public key و private key را بدست آوریم. حال کلید عمومی را به مسئول سرور می‌دهیم تا آنرا در سرور وارد کند.

\* نصب و وصل شدن به VPN - در کتابخانه ملی (سرور شرکت) از openfortivpn استفاده می‌شود. بعد از انجام تنظیمات مربوطه <sup>۱۵</sup> سعی کردم تا به آن وصل شوم. اما مشاهده کردم که خطا می‌گیرم. سپس متوجه شدم که باید دسترسی root داشته باشم. پس دستور زیر را در ترمینال <sup>۱۶</sup> وارد کردم.

```
sudo su - root
```



شکل ۶: دسترسی ریشه

و در ادامه برای این که به VPN وصل شوم، فرمان زیر را وارد کردم

<sup>۱۵</sup>Configuration

<sup>۱۶</sup>Terminal



```
openfortivpn
```

و نتیجه به صورت مقابل شد :

```
INFO: Tunnel is up and running.
```

شکل ۷: اتصال به vpn

در ادامه برای اینکه به سرور دسترسی داشته باشیم از دستور زیر استفاده کردیم:

```
ssh[servername]@[server-IP-address]
```

در ادامه برای اینکه به فایلی از سرور دسترسی داشته باشیم و آنرا دانلود کنیم، باید از scp استفاده کنیم و به این صورت دستور را وارد می‌کنیم.

```
scp [servername]:[filepath on server] [folderpath on local]
```

• نمونه‌ای از خروجی‌ای که تا به حال تولید کرده‌ام:



شکل ۸: سند هویتی (کارت ملی) تولیدشده با کد

برنامه اقدامات فردا :

• پرسش چند سوال راجع به کدها از مسئول جدید پروژه الفبا.

مشکلات پیش آمده :

در این چند وقت تنها مشکلی که داشتم کدهای برنامه بود. متأسفانه برای کدها مستندی نوشته نشده بود و کامنت <sup>۱۷</sup> گذاری مناسبی هم صورت نگرفته بود و میبایست خودم با پرسش از دیگران کار را جلو می‌بردم و این کمی کار را سخت می‌کرد.

<sup>17</sup>Comment