

بسم الله الرحمن الرحيم



دانشگاه اصفهان

دانشکده مهندسی کامپیوتر

گروه هوش مصنوعی و رباتیک

پایان نامه کارشناسی ارشد رشته‌ی مهندسی کامپیوتر

گرایش هوش مصنوعی و رباتیک

خلاصه‌سازی استخراجی متن با استفاده از رویکردهای یادگیری عمیق

استاد راهنما:

دکتر حمیدرضا برادران کاشانی

استاد مشاور:

دکتر افسانه فاطمی

دانشجو:

امیررضا صدیقین

دی‌ماه ۱۴۰۲



تعهدنامه اصالت اثر

اینجانب **امیررضا صدیقین** دانشجوی مقطع کارشناسی ارشد رشته مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیک متعهد می‌شوم که مطالب مندرج در این پایان نامه و برونادهای حاصل از آن، دستاورد پژوهشی اینجانب با اشراف و راهنمایی استاد **دکتر حمیدرضا برداران کاشانی** است و آن دسته از مطالب این پایان نامه که حاصل تحقیقات دیگران بوده نیز طبق شیوه‌نامه‌های مصوب ارجاع، مستند شده و در فهرست منابع و مآخذ این پژوهش آمده است. ضمناً اظهار می‌دارم که این پایان نامه پیش‌تر برای دریافت هیچ مدرک تحصیلی ارائه نشده است. بدیهی است دانشگاه اصفهان برای خود این حق را قائل است که در صورت احراز و اثبات هر گونه تخلف در این باره، مدرک تحصیلی اینجانب را از درجه اعتبار ساقط نماید و ضمن درج موضوع در جراید کثیر انتشار، کلیه امتیازات و حقوقی را که به موجب آن پس از دوران تحصیل، از آنها بهره‌مند گشته‌ام، از اینجانب سلب و به طرف ذی‌نفع مسترد کند.

کلیه حقوق مادی و معنوی این اثر متعلق به دانشگاه اصفهان است.

امضاء

نام و نام خانوادگی دانشجو: **امیررضا صدیقین**

نام و نام خانوادگی استاد راهنما: **دکتر حمیدرضا برداران کاشانی** امضاء



حوزه معاونت پژوهش و فناوری

دانشگاه اصفهان

دانشکده مهندسی کامپیوتر

گروه هوش مصنوعی و رباتیک

پایان نامه آقای

امیررضا صدیقین

دانشجوی رشته‌ی مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیک

خلاصه‌سازی استخراجی متن با استفاده از رویکردهای یادگیری عمیق

به عنوان بخشی از ملزومات درجه کارشناسی ارشد

در تاریخ توسط هیات داوران زیر بررسی و با درجه به تصویب نهایی رسید.

- | | | | |
|------------------------|------------------------------|-------------------------|------|
| ۱- استاد راهنمای رساله | دکتر احمد رضا برادران کاشانی | با مرتبه علمی استاد یار | امضا |
| ۲- استاد مشاور رساله | دکتر افسانه فاطمی | با مرتبه علمی استاد یار | امضا |

مهر و امضای مدیر گروه

سپاسگزاری

سپاس از اساتید بزرگوار و دوستان خوبم که در این زمینه مرایای نمودند.

سپاس از خانواده‌ی مهربانم که حامی بنده در این مسیر سخت بودند.

سپاس از تمام افرادی که در این مسیر مرایای کردند.

تقدیم به

پدر و مادر عزیزم که من را در این مسیر همراهی کرده اند.

چکیده

خلاصه‌سازی خودکار متن، یک حوزه فعال در پژوهش‌های پردازش زبان طبیعی است که با استفاده از شبکه‌های عصبی عمیق و گراف، بهبودهای چشمگیری را در این زمینه به وجود آورده است. در این تحقیق، رویکرد استخراجی برای خلاصه‌سازی متن مورد توجه قرار گرفته و از گراف‌های معنایی و نحوی برای بهبود دقت و کیفیت خلاصه‌ها بهره گرفته شده است. همچنین در این پژوهش سعی شده است تأثیر دیدگاه‌ها و روابط متفاوت در متن برای خلاصه‌سازی استخراجی متن مورد بررسی قرار گیرد.

مسئله خلاصه‌سازی به عنوان یک وظیفه مهم در پردازش زبان طبیعی، به دنبال تولید خلاصه‌های مفهومی از متون و اسناد طولانی است. در این راستا، از شبکه‌های عصبی عمیق به عنوان ابزار اصلی برای استخراج ویژگی‌های مهم از متون استفاده می‌شود. با استفاده از گراف‌های معنایی و نحوی، روابط بین کلمات و جملات و اطلاعات معنایی و ساختار بهتری می‌توان استخراج کرد. این گراف‌ها بر اساس روابط معنایی بین کلمات و ساختارهای گرامری جملات و روابط مختلف بین ماهیت‌های موجود در متن ساخته می‌شوند. همچنین در این پژوهش از یک گراف ناهمگن به منظور مدل‌سازی سند بهره برده و سعی شده است تا انواع اطلاعات با استفاده از روش‌ها و دیدگاه‌های مختلف استخراج و در خلاصه‌سازی متن مورد استفاده قرار گیرند. مدل با استفاده از داده‌های آموزشی به شیوه‌ی نظارتی آموزش می‌بیند. با استفاده از تابع خطا مبتنی بر خروجی مورد انتظار، مدل بهبود یافته و وزن‌های شبکه به‌روزرسانی می‌شوند. ارزیابی‌های انجام شده نشان‌دهنده‌ی افزایش قابل توجهی در انتقال معنا و حفظ اطلاعات کلیدی در خلاصه‌ها است. ترکیب شبکه‌های عصبی عمیق با گراف و استفاده از دیدگاه‌ها و روابط مختلف در حوزه خلاصه‌سازی، به عنوان یک رویکرد نوآورانه، نتایج موفقیت‌آمیزی را از خود نشان داده و باعث بهبود قابل توجهی در کیفیت و دقت خلاصه‌ها می‌شود. این پژوهش نشان می‌دهد که استفاده از این رویکرد باعث پیشرفت واضح در زمینه خلاصه‌سازی خودکار متن می‌شود.

کلیدواژه‌ها: ۱- خلاصه‌سازی استخراجی متن ۲- شبکه‌های عصبی عمیق ۳- گراف معنایی ۴- گراف نحوی ۵- پردازش زبان طبیعی ۶- شبکه‌های عصبی گراف ۷- خلاصه‌سازی مبتنی بر عنوان

فهرست مطالب

فصل اول بیان مسئله‌ی پژوهشی.....	۱
۱-۱- مقدمه.....	۱
۲-۱- کلیات پژوهش.....	۲
۳-۱- سوالات پژوهش.....	۶
۴-۱- اهداف پژوهش.....	۷
۵-۱- روش ارائه شده.....	۷
۶-۱- سازماندهی مطالب.....	۸
فصل دوم ادبیات موضوع.....	۹
۲-۱- مقدمه.....	۹
۲-۲- یادگیری ماشین.....	۱۰
۳-۲- یادگیری عمیق.....	۱۱
۲-۳-۱- شبکه‌ی عصبی پرسپترون چند لایه و مفاهیم اولیه یادگیری عمیق.....	۱۱
۲-۳-۲- شبکه‌های عصبی پیچشی.....	۱۵
۳-۳-۲- شبکه‌های عصبی بازگشتی.....	۱۶
۴-۳-۲- ساختار کدگذار-کدگشا.....	۱۷
۵-۳-۲- روش‌های مبتنی بر توجه و معماری ترنسفرمر.....	۱۸
۶-۳-۲- شبکه‌های عصبی گراف.....	۲۰
۴-۲- جانشانی کلمه.....	۲۳
۱-۴-۲- جانشانی کلمه Glove.....	۲۳
۲-۴-۲- مدل BERT.....	۲۴
۵-۲- معیارهای شباهت بین بردارها.....	۲۵
۶-۲- شاخص‌های ارزیابی در مسئله‌ی خلاصه‌سازی متن.....	۲۵
۱-۶-۲- معیار صحت یا Precision.....	۲۶
۲-۶-۲- معیار Recall.....	۲۶

۲۶ F-measure معیار ۳-۶-۲
۲۶ ROUGE معیار ۳-۶-۲
۲۸ جمع‌بندی ۷-۲
۲۹ فصل سوم پیشینه پژوهش
۲۹ ۱-۳ - مقدمه
۳۲ ۲-۳ - مدل‌های مبتنی بر روش‌های سنتی
۳۳ ۳-۳ - مدل‌های مبتنی بر یادگیری ماشین
۳۴ ۴-۳ - مدل‌های مبتنی بر یادگیری عمیق
۳۵ ۱-۴-۳ - مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی
۴۱ ۲-۴-۳ - مدل‌های مبتنی بر شبکه‌های عصبی گراف
۴۶ ۳-۴-۳ - روش‌های مبتنی بر عنوان
۵۱ ۵-۳ - جمع‌بندی
۵۲ فصل چهارم روش پیشنهادی
۵۲ ۱-۴ - مقدمه
۵۳ ۱-۴ - ساختار کلی معماری پیشنهادی
۵۴ ۲-۴ - ساختار SWGraph
۵۷ ۳-۴ - ساختار CNN-LSTM
۵۸ ۴-۴ - ساختار جانشانی مکانی
۵۹ ۵-۴ - ساختار KeywordGraph
۶۰ ۶-۴ - ساختار SentSyntaxGraph
۶۱ ۷-۴ - ساختار SentSemanticGraph
۶۲ ۸-۴ - ساختار DocSemanticGraph
۶۲ ۹-۴ - مدل‌های ارائه‌شده
۶۶ ۱۰-۴ - جمع‌بندی
۶۸ فصل پنجم آزمایش و نتایج

۶۸	۱-۵- مقدمه.....
۶۹	۲-۵- جزئیات پیاده سازی.....
۶۹	۳-۵- مجموعه داده.....
۷۰	۴-۵- پیش پردازش.....
۷۱	۵-۵- ابرپارامترها.....
۷۱	۶-۵- فرآیند یادگیری.....
۷۲	۷-۵- نتایج و مقایسه با روش های دیگر.....
۷۷	۸-۵- جمع بندی.....
۷۸	فصل ششم نتیجه گیری و کارهای آینده.....
۷۸	۱-۶- مقدمه.....
۷۸	۲-۶- نتیجه گیری.....
۷۹	۳-۶- کارهای آینده.....

فهرست شکل‌ها

- شکل ۱: نمای کلی از مراحل یک سیستم خلاصه‌سازی خودکار متن ۳
- شکل ۲: نمونه‌ای از خلاصه‌سازی خودکار متن با رویکرد استخراجی ۴
- شکل ۳: نمای کلی از معماری سیستم‌های خلاصه‌سازی خودکار متن با رویکرد استخراجی ... ۴
- شکل ۴: ساختار یک شبکه عصبی که دارای سه بخش اصلی لایه‌ی ورودی، لایه مخفی و لایه‌ی خروجی می‌باشد..... ۱۲
- شکل ۵: نحوه‌ی محاسبه‌ی خروجی هر نرون در شبکه‌ها عصبی. هر نرون از چندین ورودی و وزن به همراه بایاس (مقدار وزن با ورودی ۱) تشکیل شده و با استفاده از تابع غیر خطی f ویژگی غیر خطی به خروجی می‌دهد..... ۱۲
- شکل ۶: برخی توابع فعالسازی معروف ۱۳
- شکل ۷: ساختار residual باعث جلوگیری از گم شدن گرادیان می‌کند. ۱۴
- شکل ۸ : مراحل بدست آوردن خروجی از اعمال عملیات پیچش یک فیلتر بر روی یک ماتریس ورودی ۱۵
- شکل ۹: نحوه‌ی حرکت کرنل پیچشی یک بعدی..... ۱۶
- شکل ۱۰: نحوه‌ی عملکرد شبکه‌های عصبی بازگشتی ۱۷
- شکل ۱۱: شبکه‌های GRU و LSTM و تفاوت آن‌ها با شبکه‌های RNN اولیه. ۱۷
- شکل ۱۲: ساختار کدگذار-کدگشا ۱۸
- شکل ۱۳: نمایش مکانیزم توجه برای بدست آوردن نمایش برای یک دنباله ۱۹
- شکل ۱۴: گراف‌ها می‌توانند وزن دار یا بی وزن، جهت دار یا بی جهت، دوردار یا بی دور باشند .. ۲۱
- شکل ۱۵: نحوه‌ی انتقال پیام برای یک گره مشخص در شبکه‌ی GCN..... ۲۲
- شکل ۱۶: ادغام گره‌های گراف و تولید گراف کوچکتر با استفاده از لایه‌ی ادغام گراف ۲۳
- شکل ۱۷: جانشانی کلمه در مدل BERT با استفاده از سه جانشانی توکی، بخش و موقعیت تولید می‌شود. ۲۴
- شکل ۱۸: لایه‌های مدل SummaRuNNer..... ۳۶
- شکل ۱۹: معماری کلی LatentExtSum..... ۳۷

شکل ۲۰: استخراج ویژگی‌های جملات به صورت سلسله مراتبی در دو سطح جمله و سند	۳۸
شکل ۲۱: مدل BERT اصلی و BERT برای خلاصه‌سازی متن.	۴۰
شکل ۲۲: کدگذاری سلسله مراتبی جملات در روش HIBERT	۴۱
شکل ۲۳: معماری کلی از شبکه‌ی DiscoBert	۴۲
شکل ۲۴: معماری کلی و ساخت گراف در مدل HSG	۴۴
شکل ۲۵: شمای کلی گراف در مدل HDSG	۴۴
شکل ۲۶: معماری شبکه Multi-Gras	۴۵
شکل ۲۷: معماری مدل TopicGraphSum	۴۷
شکل ۲۸: نمای کلی از روش ارائه شده توسط بلوال و همکاران و استفاده از گراف و مجموعه کلیدواژه‌ها یا عناوین برای خلاصه‌سازی متن.	۴۹
شکل ۲۹: معماری کلی روش DeepSum	۵۰
شکل ۳۰: روند کلی معماری خلاصه‌سازی خودکار متن با رویکرد استخراجی.	۵۳
شکل ۳۱: نمای کلی معماری SWGraph برای بدست آوردن بردار ویژگی‌های جملات.	۵۵
شکل ۳۲: نمونه‌ای گراف SWGraph برای یک متن.	۵۶
شکل ۳۳: شمای کلی ساختار CNN-LSTM در مدل پیشنهادی.	۵۷
شکل ۳۴: میزان وقوع حضور جملات در خلاصه با جایگاه مکانی متفاوت در سند داخل دو مجموعه داده CNN/DailyMail و arXive	۵۸
شکل ۳۵: کلیدواژه‌های استخراجی از یک نمونه متن و ساخت گراف متناظر با آن.	۵۹
شکل ۳۶: مثالی از گراف نحوی تولید شده برای یک جمله.	۶۰
شکل ۳۷: معماری ساختار SentSyntaxGraph و SentSemanticGraph	۶۱
شکل ۳۸: مدل اولیه با استفاده از ساختار SWGraph	۶۳
شکل ۳۹: ترکیب ساختار SentSemanticGrph با SWGraph به صورت توالی	۶۵
شکل ۴۰: ترکیب ساختار SentSemanticGrph با SWGraph به صورت موازی	۶۵
شکل ۴۱: ترکیب مدل SWGraph و SentSemanticGraph با دیگر مدل‌های استخراج کننده‌ی بردار ویژگی جملات.	۶۵

فهرست جداول

جدول ۱: چند تابع خطا معروف در شبکه‌های عصبی.....	۱۳
جدول ۲: تعداد سندها در مجموعه داده CNN/DailyMail.....	۷۰
جدول ۳: میانگین تعداد کلمات در هر سند.....	۷۰
جدول ۴: میانگین تعداد جملات در هر سند.....	۷۰
جدول ۵: نتایج بدست آمده از ترکیب ساختارهای مختلف ارائه شده.....	۷۴
جدول ۶: نتایج بدست آمده آزمایش‌های مختلف با اعمال تغییرات در بخش‌های مختلف مدل	
SWGraph + SentSemanticGraph.....	۷۵
جدول ۷: نتایج مدل‌های معروف در حوزه‌ی خلاصه‌سازی خودکار متن با رویکرد استخراجی و	
مقایسه آن‌ها با مدل پیشنهادی.....	۷۶

فصل اول

بیان مسئله‌ی پژوهشی

۱-۱- مقدمه

در دهه‌های گذشته، نحوه‌ی ذخیره و توزیع اطلاعات به صورت چشمگیری تغییر پیدا کرده است. در حالی که کتابخانه‌ها مسئولیت نگهداشت تعداد زیادی کتاب چاپی بوده است، امروزه بخش زیادی از محتوای متنی به صورت الکترونیکی در دسترس هستند [1]. در جهان امروز، روزانه منابع متنی گسترده‌ای با سرعت بالا در حال تولید هستند، این منابع در تنوع و حجم بالا، مخزن غنی از اطلاعات می‌باشند [2]. با این حال، این توسعه یک چالش را به همراه دارد. مقادیر متنی موجود برای هر موضوع معین، معمولاً آنقدر زیاد است که یک فرد بتواند در یک زمان معقول آن را پردازش کند. به عنوان مثال، نسخه‌ی انگلیسی ویکی‌پدیا در سپتامبر ۲۰۱۸، حاوی ۶ میلیون مقاله و سند بود. گوگل در قالب یک پروژه^۱، بیش از ۲۵ میلیون نسخه کتاب را تا سال ۲۰۱۵ دیجیتالی کرده بود. همچنین تعداد کل صفحات شاخص گذاری شده، در سپتامبر ۲۰۱۸ برابر ۴/۴ میلیارد برآورد شده است [3]. فرآیند خلاصه‌سازی دستی متن بدون شک راهی موثر برای حفظ معنا و مفهوم متن است. با این حال،

¹ Google Books

خلاصه‌سازی دستی متن، یک فعالیت وقت‌گیر و پرهزینه می‌باشد [4]. به همین منظور، مسئله‌ای در حوزه‌ی پردازش زبان‌های طبیعی^۱، به عنوان خلاصه‌سازی خودکار متن^۲ به وجود آمده است.

خلاصه‌سازی متن یک فرآیند مهم در پردازش زبان طبیعی است که در آن سعی می‌شود اطلاعات کلیدی و مهم متن اصلی به طور خلاصه و مختصری باقی بماند. این فرآیند به افزایش دسترسی به اطلاعات برای مخاطبین، صرفه‌جویی در زمان و انرژی، و همچنین ایجاد یک نمای کلی از متن‌های بزرگتر کمک می‌کند. یکی از رویکرد های خلاصه‌سازی خودکار متن، خلاصه‌سازی استخراجی متن^۳ می‌باشد که در آن جملات مهم متن استخراج می‌شوند.

شبکه‌های عصبی^۴ به علت قدرت پردازشی و توانایی فهم و تحلیل داده‌های پیچیده از اهمیت بالایی در خلاصه‌سازی متن برخوردار هستند. با استفاده از شبکه‌های عصبی می‌توان اطلاعاتی را از متن‌ها استخراج کرد و یا جملات جدیدی را برای تشکیل خلاصه ایجاد کرد. این شبکه‌ها می‌توانند از مدل‌های زبانی^۵ پیشرفته مانند ترنسفرمر^۶ ها استفاده کنند که قدرت بسیار بالایی در فهم زبان و تولید محتوا دارند [5].

۱-۲- کلیات پژوهش

هدف سیستم‌های خلاصه‌سازی خودکار متن، ساخت خلاصه‌هایی حاوی اطلاعات مهم و مرتبط با سند مربوطه، در زمان کوتاه و هزینه‌ی کم می‌باشند [3]. عمده‌ی کاربردهای این سیستم‌ها در سیستم‌های بازیابی اطلاعات^۷، استخراج اطلاعات^۸ و جواب‌دهی به سوالات^۹ است. برای مثال در سیستم‌های بازیابی اطلاعات، از خلاصه‌سازی خودکار متن برای تقویت موتورهای جستجو استفاده می‌شود. منابع متنی می‌تواند در حوزه‌های مختلفی باشند، برای همین می‌توان کاربردهای متنوعی برای سیستم‌های خلاصه‌سازی در نظر گرفت که خلاصه‌سازی اخبار، نظرات، کتاب، داستان، ایمیل، مقالات علمی، توییتهای و اطلاعات شبکه‌های اجتماعی، از موارد آنها می‌باشند. همانطور که در شکل ۱ آمده است، یک سیستم خلاصه‌سازی خودکار متن دارای مراحل پیش‌پردازش^{۱۰}، پردازش و

¹ Natural language processing

² Automatic text summarization

³ Extractive text summarization

⁴ Neural networks

⁵ Language models

⁶ Transformer

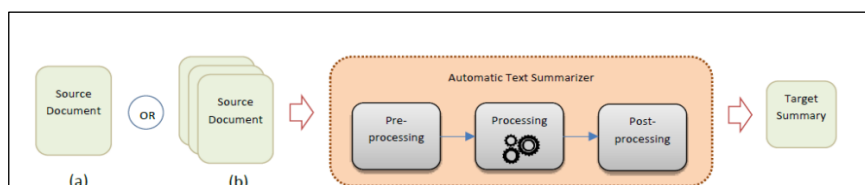
⁷ Information retrieval

⁸ Information extraction

⁹ Question answering

¹⁰ Preprocessing

پس‌پردازش^۱ متن است. در قسمت پیش‌پردازش متن، یک نمایش ساختار یافته از متن به دست می‌آید [2]. از عملیات موجود در این بخش می‌توان به برچسب‌گذاری POS^۲، جداسازی کلمات توقف^۳، هرس کردن^۴ و محاسبه‌ی ویژگی‌ها با رویکردهای مختلف، اشاره کرد. در قسمت پردازش، یک رویکرد خلاصه‌سازی در نظر گرفته می‌شود و با توجه به آن خلاصه به دست می‌آید [4]. در پس‌پردازش نیز برخی مشکلات خلاصه‌ی به دست آمده، برای مثال مشکل عدم تشخیص مرجع ضمائر یا مرجع زمان‌های نسبی، بر طرف می‌شود [2]. انتخاب ویژگی‌های مناسب یکی از مهم‌ترین گام‌های موجود در تولید یک مدل خلاصه‌سازی می‌باشد. شباهت کلمات یک جمله با عنوان^۵، موقعیت جملات در متن، ویژگی‌های TF-IDF^۶، طول جملات، کلمات موضوعی (کلمات مرتبط با دامنه و دارای میزان تکرار زیاد)، اسامی خاص، شباهت بین جمله‌ای و اعداد، ویژگی‌های مهمی هستند که در یک سیستم خلاصه‌سازی متن مورد توجه قرار می‌گیرند [4].



شکل ۱: نمای کلی از مراحل یک سیستم خلاصه‌سازی خودکار متن [4].

سیستم‌های خلاصه‌سازی را از زاویه‌های متنوع دسته‌بندی کرد. از لحاظ تعداد سند، این سیستم‌ها به دو دسته‌ی تک‌سندی و چندسندی تقسیم می‌شوند. همچنین برپایه‌ی رویکردهای موجود در این نوع سیستم‌ها، سه رویکرد استخراجی^۷، انتزاعی^۸ یا اسنتتاجی و ترکیبی^۹ وجود دارد. در رویکرد استخراجی، جملات خلاصه عیناً از جملات متن مرجع استخراج می‌شوند. در رویکرد انتزاعی، سعی در درک متن مرجع و ساخت خلاصه‌ی منسجم نزدیک به خلاصه‌های انسانی است. در رویکرد ترکیبی نیز، ابتدا با رویکرد استخراجی، جملات از متن ورودی استخراج، سپس با استفاده از این جملات، یک خلاصه‌ی منسجم تولید می‌شود. در ادامه به بررسی دقیق‌تر رویکردهای حوزه‌ی خلاصه‌سازی متن، پرداخته خواهد شد [2]. شکل ۲ نمونه‌ای از خلاصه‌سازی متن با رویکرد

¹ Postprocessing

² Part of speech tagging

³ Stop words Filtering

⁴ Stemming

⁵ Title similarity

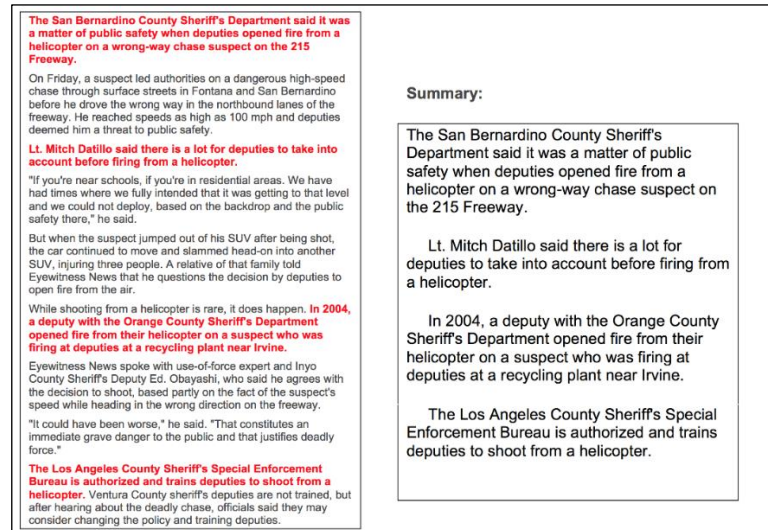
⁶ Term frequency – inverse term frequency

⁷ Extractive

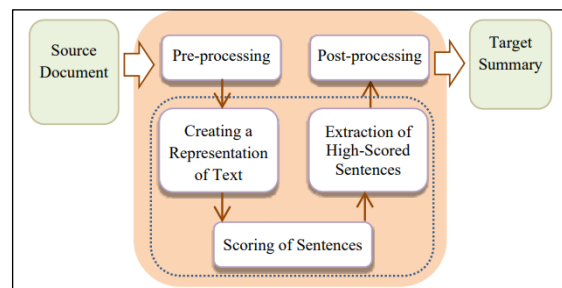
⁸ abstractive

⁹ hybrid

استخراجی است. همچنین شکل ۳ نمای کلی از مراحل یک سیستم خلاصه‌سازی خودکار متن با رویکرد استخراجی را نشان می‌دهد. در این مقاله به رویکرد استخراجی در خلاصه‌سازی متن‌های تک‌سندی پرداخته شده است.



شکل ۲: نمونه‌ای از خلاصه‌سازی خودکار متن با رویکرد استخراجی [6].



شکل ۳: نمای کلی از معماری سیستم‌های خلاصه‌سازی خودکار متن با رویکرد استخراجی [2].

برای ساخت یک سیستم خلاصه‌سازی، چالش‌های زیادی وجود دارد. شناسایی بخش‌های مهم متن، خلاصه‌سازی سندهای بزرگ مثل کتاب، خلاصه‌سازی هم‌زمان چند سند در قالب یک سیستم، ارزیابی سیستم و ساخت خلاصه شبیه به خلاصه‌های انسان از جمله این چالش‌ها هستند [2]. در ادامه به برخی از چالش‌های اصلی در رویکرد استخراجی پرداخته خواهد شد.

• مشکل افزونگی^۱ [3]

¹ Problem of redundancy

افزونگی باعث کاهش کیفیت یک خلاصه می‌شود. وظیفه خلاصه‌ی یک متن علاوه بر دور انداختن مطالب زائد، انتقال محتوای اصلی متن نیز می‌باشد. اگر مطالب تکراری را حذف کنیم، می‌توانیم اطلاعات مهم بیشتری در خلاصه جای دهیم. یکی از روش‌های جلوگیری از مشکل افزونگی، استفاده از اندازه‌گیری‌های شباهت بین جملات است.

- مشکل بی‌ربط بودن^۱ [3]

هدف اصلی یک سیستم خلاصه‌سازی متن، استخراج مطالب مرتبط با سند در یک نمای سریع و کلی، است. برخی از ویژگی‌های مورد استفاده در خلاصه‌سازی، ممکن است تمایل به ایجاد محتوای فرعی یا نامربوط در خلاصه داشته باشند. از این رو، بسیار مهم است که بدانیم کدام ویژگی‌ها و با چه میزان نسبت توجه به آن‌ها، برای ایجاد خلاصه با کیفیت مناسب هستند.

- مشکل پوشش ندادن تمام موضوعات اصلی^۲ [3]

پوشش تمام موضوعات اصلی سند در خلاصه، یک موضوع اساسی در سیستم‌های خلاصه‌سازی است. یک خلاصه عمومی خوب باید اطلاعات مربوط به هر جنبه از سند را ذکر کند. بیشتر تکنیک‌های خلاصه‌سازی کنونی، تضمین پوشش تمام موضوعات اصلی را به کاربر نمی‌دهند. این مشکل بیشتر در خلاصه‌سازی چند سندی به دلیل تنوع و پراکندگی بیشتر موضوعات، قابل مشاهده هستند.

- مشکل ناخوانایی و انسجام کم محتوا^۳ [3]

یک خلاصه خوب باید خوانا و منسجم باشد. منظور از خواندنی و منسجم این است که مطالب خلاصه باید از نظر مفهومی با یکدیگر مرتبط و پیوستگی داشته باشند. بیشتر روش‌های استخراجی، فاقد این ویژگی هستند.

- چالش در نظر گرفتن روابط بین جمله‌ای با فاصله زیاد از هم^۴ [7]

جملات موجود در یک سند، هر چند با فاصله از هم ممکن است روابط و وابستگی‌های زیادی با هم داشته باشند. بیشتر مدل‌های پیشنهادی ارائه شده، توجه زیادی به روابط بین جمله‌ای ندارند. مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی، تا حدی این مورد را در نظر

¹ Problem of irrelevancy

² Problem of loss of coverage

³ Problem of non-readability and less cohesive content

⁴ Long-distance relationships

می‌گیرند. با این حال مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی، نمی‌توانند به‌خوبی روابط بین جملات با فاصله‌ی زیاد از هم را در نظر بگیرند.

تأثیر شبکه‌های عصبی بر روی خلاصه‌سازی خودکار متن بسیار قابل توجه است. این شبکه‌ها با بهره‌گیری از معماری‌های عمیق و مدل‌های توجه، قدرتمندترین ابزارها برای انجام خلاصه‌سازی متن هستند. با پیشرفت تکنولوژی شبکه‌های عصبی و استفاده از مجموعه داده‌های بزرگ، می‌توان به دقت و کیفیت بالاتری در خلاصه‌های تولید شده دست یافت. همچنین، توانایی‌های ترنسفورمر و شبکه‌های بازگشتی^۱ (RNN) در فهم ارتباطات طولانی‌تر متن و اجزای آن، موجب بهبود چشم‌گیری در تولید خلاصه‌های معنادار و مرتبط با محتوای متن اصلی شده است. از طرفی، شبکه‌های عصبی می‌توانند با تعمیم‌دهی و یادگیری از داده‌های متنوع و چندزبانه، بهبود مسائل ترجمه و خلاصه‌سازی متن را در زبان‌های مختلف ممکن سازند. به‌طور خلاصه، تحولات در حوزه شبکه‌های عصبی، از اهمیت بسیاری برای ارتقاء و بهبود روش‌های خلاصه‌سازی خودکار متن برخوردار است [8]. همچنین تأثیر گراف و شبکه‌های عصبی گراف بر روی خلاصه‌سازی خودکار متن از جمله موضوعات پرتعداد در حوزه پردازش زبان طبیعی است. در این رویکرد، اطلاعات متنی به صورت گراف ترسیم می‌شوند که اجزا و ارتباطات میان واژه‌ها را نشان می‌دهد. این گراف‌ها می‌توانند با استفاده از روش‌های مبتنی بر شبکه‌های عصبی گراف، مورد استفاده قرار گیرند تا خلاصه‌های متنی با کیفیت و منطبق با محتوای اصلی ایجاد شوند. شبکه‌های عصبی گراف با توجه به ارتباطات بین واژه‌ها و ساختار معنایی متن، می‌توانند بهبود قابل توجهی در فهم و تولید خلاصه‌های معنادار داشته باشند [9], [10].

۱-۳- سوالات پژوهش

در این پژوهش، به سؤالاتی در زمینه‌ی خلاصه‌سازی خودکار متن با رویکرد استخراجی پاسخ داده شده است. بخش نتایج و پیشینه پژوهش به ارزیابی تأثیر شبکه‌های عصبی، به ویژه شبکه‌های عصبی گراف، و نقش انواع ساختارهای گراف و دیدگاه‌های متفاوت در مسئله‌ی خلاصه‌سازی متن پرداخته است. این بخش‌ها همچنین به بررسی موارد دیگری نظیر ویژگی‌ها و پیش‌پردازش‌های متناسب با مسئله‌ی خلاصه‌سازی و چگونگی استخراج ویژگی‌های مناسب برای این مسئله می‌پردازند. سؤالاتی که در طی این پژوهش به آن‌ها پاسخ داده شده و مورد بررسی قرار گرفته است به شرح زیر می‌باشد.

¹ Recurrent Neural Network

۱- معماری‌های متفاوت شبکه‌های عصبی عمیق، به خصوص شبکه‌های عصبی گراف، تا چه میزان می‌توانند در زمینه‌های خلاصه‌سازی خودکار متن با رویکرد استخراجی موفق ظاهر شوند؟

۲- استفاده از روش‌های مختلف ساخت گراف تا چه میزان بهبود برای مسئله‌ی خلاصه‌سازی متن را می‌تواند حاصل کند؟

۳- چه ویژگی‌هایی از متن و به چه میزان در بهبود خلاصه‌سازی متن کمک می‌کنند؟

۴- موجودیت‌های یک متن در مسئله‌ی خلاصه‌سازی چه مواردی هستند و به چه میزان، در بهبود خلاصه‌سازی متن کمک می‌کنند؟

۴-۱- اهداف پژوهش

در زمینه‌ی خلاصه‌سازی متن، با وجود تعداد زیادی از روابط و متغیرها، توسعه‌ی مدل‌های مؤثر و مناسب برای مدل‌سازی این روابط امری حیاتی است. هدف این پژوهش، ایجاد یک مدل پیشرفته با استفاده از شبکه‌های عصبی به عنوان ابزار اصلی مدل‌سازی برای خلاصه‌سازی خودکار متن با رویکرد استخراجی می‌باشد. در این پژوهش، معماری‌های مختلف شبکه عصبی، تأثیر ساختارهای گرافی بر نتایج و ویژگی‌های بهینه برای این مسئله مورد بررسی قرار گرفته‌اند. همچنین در این پژوهش، ارتقاء کیفیت خلاصه‌ها و افزایش دقت انتخاب جملات مهم متن با استفاده از شبکه‌های عصبی، به خصوص شبکه‌های عصبی گراف و تجزیه و تحلیل نقش و تأثیر ساختارهای گرافی مختلف بر مسئله‌ی خلاصه‌سازی متن مورد توجه قرار گرفته است.

در این تحقیق، داده‌های آموزش به صورت جفت متون و خلاصه‌های متناظر آنها فراهم شده‌اند. با بهره‌گیری از یک روش برچسب‌گذاری و ارتباط خلاصه با جملات متن ورودی، مسئله‌ی خلاصه‌سازی متن به یک مسئله‌ی دسته‌بندی دودویی تبدیل گردیده و برای آموزش و ارزیابی مدل‌ها از رویکرد نظارتی بهره گرفته شده است.

انتظار می‌رود که با اعمال بهبودها و نوآوری‌های پیشنهادی، کیفیت و دقت خلاصه‌های تولیدی توسط مدل‌های عصبی افزایش یابد. همچنین، نتایج بررسی تأثیر ساختارهای گراف و دیدگاه‌های متفاوت در تولید خلاصه‌ها به وضوح مشخص خواهد شد.

۵-۱- روش ارائه شده

در این پایان‌نامه، یک سیستم خودکار خلاصه‌سازی متن با رویکرد استخراجی معرفی می‌شود که از ویژگی‌های جملات در سطوح و دیدگاه‌های متنوع برای بهبود در این مسئله استفاده می‌کند. این پژوهش از بخش‌های مختلف برای استخراج ویژگی‌های جملات در یک متن بهره می‌برد. از لایه‌های

مختلف شبکه‌های عصبی برای استخراج ویژگی‌های دنباله‌ای، از گراف ناهمگن با گره‌های جمله و کلمات برای استخراج ویژگی‌های زمینه‌ای در سطح سند، از گراف معنایی برای استخراج ویژگی‌های معنایی، و از گراف نحوی برای استخراج ویژگی‌های نحوی در جملات بهره گرفته شده است. همچنین، به منظور بهبود استخراج ویژگی‌ها در سطح سند و بهره‌گیری از کلیدواژه‌های یک متن، از دو نوع گراف کلیدواژه و گراف معنایی در سطح متن استفاده شده است.

با اجرای آزمایش‌های گسترده بر روی مدل‌های ارائه شده بر روی مجموعه داده CNN/DailyMail [11] و انجام مقایسه با سایر مدل‌ها، تأثیر دیدگاه‌های مختلف مورد بررسی قرار گرفته و بهبودهای قابل توجهی کسب شده است.

۱-۶- سازماندهی مطالب

مطالبی که در فصول آینده آمده است به قرار زیر است:

- فصل دوم به ادبیات موضوع و توضیح مفاهیم پایه برای استفاده در این پژوهش پرداخته است. در این بخش مواردی همچون مفاهیم شبکه‌ی عصبی، سنج‌ها و معیارهای مناسب در مسئله‌ی خلاصه‌سازی متن و همچنین مباحث ریاضی و گرافی در این زمینه وجود دارند.
- در فصل سوم کارهای پیشین در این حوزه و چالش‌های آن‌ها بررسی شده است. برخی تحقیقات و مدل‌های مبتنی بر شبکه‌ی عصبی که در حوزه‌ی خلاصه‌سازی خودکار متن موجود است، در این بخش ذکر شده است.
- فصل چهارم شامل توضیح روش ارائه شده و بررسی جوانب مختلف آن می‌باشد. در این بخش برخی روش‌های آزمایش شده به همراه بررسی آن‌ها وجود دارند.
- در فصل پنجم نتایج آزمایش‌های روش‌های ارائه شده بر روی مجموعه داده مختلف بررسی شده و نتایج مورد تحلیل و بررسی قرار گرفته است.
- در نهایت در فصل ششم نتیجه‌گیری نهایی و جمع‌بندی پژوهش ارائه و پیشنهاداتی برای ادامه‌ی پژوهش داده شده است.

فصل دوم

ادبیات موضوع

۱-۲- مقدمه

در حل مسائل خلاصه‌سازی متن، از ساختارها و روش‌های مختلفی استفاده می‌شود. در اکثر روش‌های نوین، یادگیری عمیق و ساختارهای آن به عنوان چهارچوب اصلی مورد توجه قرار می‌گیرند. شبکه‌های عصبی پیچشی (CNN)^۱ و شبکه‌های عصبی بازگشتی (RNN)^۲ از جمله ابزارهای مهم در زمینه‌های پردازش زبان طبیعی^۳ و متن‌کاوی^۴ به شمار می‌روند. همچنین، با ظهور شبکه‌های عصبی گراف^۵، تحولات چشمگیری در حوزه یادگیری عمیق^۶ رخ داده است.

¹ Convolutional neural networks

² Recurrent neural networks

³ Natural language processing (NLP)

⁴ Text mining

⁵ Graph neural networks

⁶ Deep learning

در این فصل، به بررسی برخی از ساختارهای داده‌ای اساسی، مفاهیم یادگیری ماشین، و نیز شبکه‌های عصبی و برخی تکنیک‌های آن ر پردازش زبان طبیعی پرداخته شده است. همچنین، معیارهای ارزیابی مدل‌ها در زمینه خلاصه‌سازی متن مورد بحث قرار گرفته‌اند. این تشریحات به ما کمک خواهد کرد تا در فهم بهتری از تکنیک‌ها و مفاهیم مرتبط با حوزه پژوهشمان به دست آید.

۲-۲- یادگیری ماشین^۱

یادگیری ماشین به صورت عام به دسته‌ای از الگوریتم‌ها و مدل‌های ریاضی اطلاق می‌شود که از داده‌ها به صورت خودکار الگوها و ارتباطات را استخراج کرده و از آنها برای اتخاذ تصمیمات در آینده استفاده می‌کنند. این زمینه را می‌توان به سه دسته‌ی یادگیری نظارت‌شده، یادگیری بدون نظارت و یادگیری نیمه نظارتی تقسیم کرد. در یادگیری نظارت‌شده، مدل‌ها با استفاده از داده‌های برچسب‌خورده آموزش می‌بینند و سپس برای پیش‌بینی برچسب یا مقدار جدید از این دسته داده‌ها استفاده می‌شود. در یادگیری بدون نظارت، مدل بدون داده‌های برچسب‌خورده آموزش می‌بیند و به دنبال الگوها و ساختارهای نهفته در داده‌ها می‌گردد. همچنین یادگیری نیمه نظارت‌شده، ترکیبی از دو حالت قبلی است که هم داده‌های برچسب‌خورده و هم بدون برچسب را برای آموزش مدل استفاده می‌کند. معروف‌ترین مسائل در حوزه‌ی یادگیری ماشین نیز، مسائل دسته‌بندی، خوشه‌بندی و رگرسیون می‌باشند. هدف از مسائل دسته‌بندی، پیش‌بینی برچسب یا دسته برای داده‌ی ورودی و هدف از مسائل خوشه‌بندی، گروه‌بندی داده‌ها بدون دانستن برچسب‌های آنها و بر اساس ویژگی‌های مشترک آنها است. همچنین مسائل رگرسیون، سعی بر پیش‌بینی مقدار یک عدد یا متغیر پیوسته دارند. [12], [13]. در دسته‌بندی با یادگیری ماشین، می‌توان از روش‌های متعددی برای ایجاد مدل طبقه‌بندی استفاده کرد. یکی از روش‌های معروف این زمینه الگوریتم SVM^۲ است. این روش با تعیین یک فضای جداکننده بین دسته‌های مختلف، سعی در بهینه‌سازی فاصله‌ی حاشیه بین داده‌های مختلف دارد [14]. از الگوریتم‌های مهم در حوزه‌ی دسته‌بندی می‌توان به الگوریتم k نزدیک‌ترین همسایه (KNN)^۳، درخت تصمیم^۴، جنگل تصادفی^۵ و Naïve Bayes نام برد. همچنین می‌توان از روش‌های خوشه‌بندی مثل الگوریتم k -means به منظور پیدا کردن k خوشه بر اساس معیارهای شباهت بین نقاط، برای خوشه‌بندی و شباهت سنجی بین اسناد استفاده کرد.

¹ Machine learning

² Support vector machines

³ k-nearest neighbors

⁴ Decision tree

⁵ Random forest

یادگیری تقویتی یکی دیگر از رویکردهای یادگیری ماشین می‌باشد. در این رویکرد، مدل‌ها با یک محیط تعاملی در ارتباط هستند و از بازخوردها برای بهبود عملکرد خود استفاده می‌کنند. این روش مختص به مواردی است که عملکرد مدل مستقیماً به دنباله اقداماتی که انجام می‌دهد، وابسته است. برای مثال، در پردازش زبان طبیعی، مسائلی مانند ترجمه ماشینی یا چت‌بات‌ها از یادگیری تقویتی بهره می‌برند. مدل با انجام اقدامات مختلف، مانند تولید جملات یا پاسخ به سوالات، از تجربیات خود یاد می‌گیرد. سپس از بازخوردها و پاداش‌ها برای بهبود کارایی و تصحیح عملکرد استفاده می‌کند [15].

۲-۳- یادگیری عمیق

یادگیری عمیق یک حوزه پویا و حیاتی در علم داده^۱ و یادگیری ماشین است که با استفاده از شبکه‌های عصبی عمیق سعی دارد نمایش‌های پیچیده داده‌ها به منظور استخراج ویژگی‌های نهان بدست آورد. شبکه‌های عصبی عمیق به عنوان مدل‌های محاسباتی با تعداد لایه‌های بسیار زیاد شناخته می‌شوند. این روش‌ها با الهام از ساختار مغز انسان و اقتباس از عملکرد نورون‌ها، مفاهیمی چون یادگیری نمایش‌ها، تشخیص الگوها و پیش‌بینی‌های دقیق را فراهم می‌کنند. شبکه‌های عصبی عمیق به عنوان ابزاری کارآمد در بسیاری از حوزه‌ها از جمله پردازش تصویر، پردازش زبان‌های طبیعی، پردازش صوت و حتی بازیابی اطلاعات، به کار گرفته می‌شوند [13].

۲-۳-۱- شبکه‌ی عصبی پرسپترون چند لایه^۲ و مفاهیم اولیه یادگیری عمیق

شبکه عصبی چندلایه یکی از مدل‌های اولیه و مهم در زمینه یادگیری عمیق است. این شبکه به عنوان یک نمونه از شبکه‌های عصبی عمیق، از چند لایه از نورون‌ها (واحدهای پردازشگر) تشکیل شده است. همانطور که در شکل ۴ آمده است، در این نوع شبکه‌ها، از سه بخش اصلی به نام لایه‌ی ورودی^۳، لایه‌های پنهان^۴، و لایه خروجی^۵ تشکیل شده است. در لایه ورودی، ویژگی‌های ورودی به شبکه وارد می‌شوند. لایه‌های پنهان با انجام عملیات‌های خطی و توابع غیرخطی^۶ یا فعالسازی^۱ بر

¹ Data science

² Multi-layer perceptron (MLP)

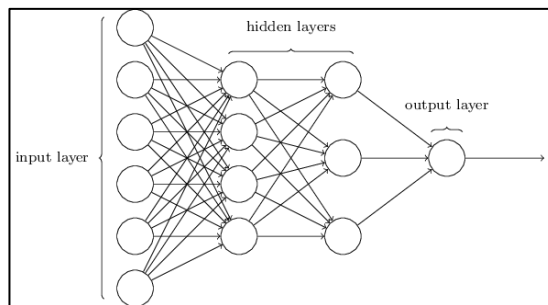
³ Input layer

⁴ Hidden layers

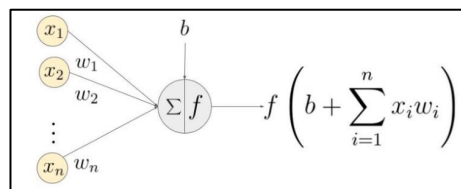
⁵ Output layer

⁶ Non-linear

روی ورودی‌ها، اطلاعات را انتزاع می‌کنند و ویژگی‌های پنهان را استخراج می‌کنند. در نهایت، لایه خروجی نتایج پایانی را تولید می‌کند. این شبکه‌ها از یک معماری کاملاً متصل² هستند، به این معنی که هر نورون در یک لایه با تمام نورون‌های لایه قبل و بعد از خود ارتباط دارد. این ویژگی باعث تغییر ابعاد داده و استخراج ویژگی‌های پیچیده‌تر می‌شود. برخی مواقع داده‌ها به صورت خطی³ تفکیک پذیر نیستند، به همین منظور در برخی از لایه‌ها از توابع غیرخطی یا فعالسازی استفاده می‌کنند. همانطور که در شکل ۵ نحوه‌ی محاسبات درون هر نورون به نمایش گذاشته شده، ورودی هر نورون خروجی لایه‌های قبل یا لایه‌ی ورودی شبکه‌ی عصبی می‌باشد. همچنین متناظر با هر مؤلفه ورودی، وزنی در آن ضرب می‌شود که وزن‌ها مقادیر قابل آموزش می‌باشند. در هر نورون یک ورودی با مقدار ۱ به همراه وزن آن (بایاس) نیز اضافه می‌شود. در آخر تابع فعالسازی بر روی مجموع ضرب ورودی با وزن‌های متناظر اعمال می‌شود. در شکل ۶ برخی توابع فعالسازی متداول، معرفی شده‌اند.



شکل ۴: ساختار یک شبکه عصبی که دارای سه بخش اصلی لایه‌ی ورودی، لایه مخفی و لایه‌ی خروجی می‌باشد [16].

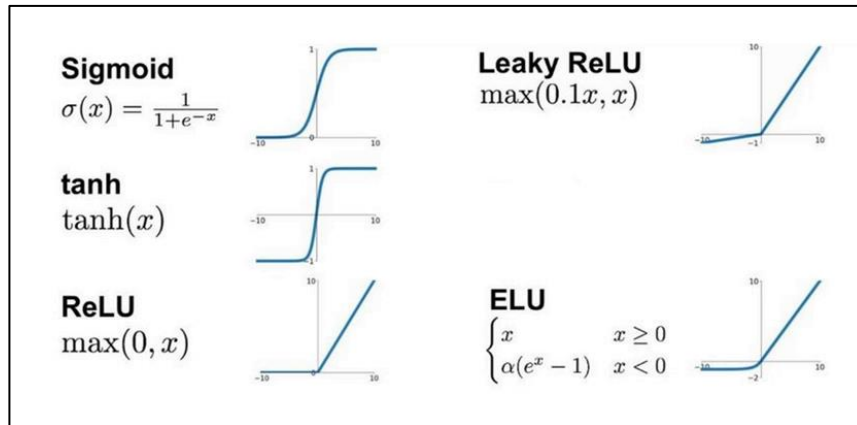


شکل ۵: نحوه‌ی محاسبه‌ی خروجی هر نورون در شبکه‌ها عصبی. هر نورون از چندین ورودی و وزن به همراه بایاس (مقدار وزن یا ورودی ۱) تشکیل شده و با استفاده از تابع غیرخطی f ویژگی غیرخطی به خروجی می‌دهد [17].

¹ Activation functions

² Fully connected

³ Linear



شکل ۶: برخی توابع فعال‌سازی معروف [18]

متناسب با نوع مسئله و الگوریتم‌های بهینه‌سازی، فرایند آموزش شبکه‌ی عصبی شامل دو عنصر مهم تابع خطا^۱ و الگوریتم بهینه‌سازی^۲ می‌باشد. در هر دور از آموزش، داده‌ها به شبکه‌ی عصبی ارسال شده و متناسب با نوع مسئله، تابع خطا مقادیر خطا را محاسبه می‌کند و با استفاده از الگوریتم‌های بهینه‌سازی که معمولاً با محاسبه‌ی مشتق خطا و قابلیت زنجیره‌ای آن همراه است، مقادیر وزن‌ها را بروزرسانی می‌کند و سعی می‌کند مقادیر خطا را به حداقل حالت خود برساند [13]. از الگوریتم‌های بهینه‌سازی معروف می‌توان به الگوریتم‌های SGD، Momentum، Adagrad، Adadelta، RMSProp و Adam اشاره کرد [19]. همچنین جدول ۱ برخی توابع خطا را معرفی می‌کند.

جدول ۱: چند تابع خطا معروف در شبکه‌های عصبی [20].

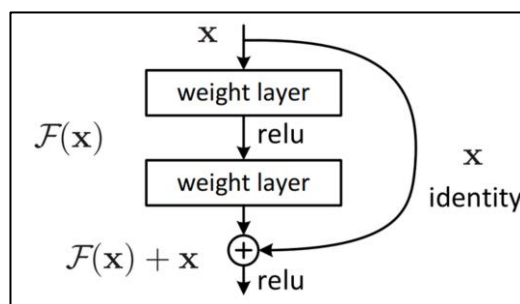
نام تابع	رابطه‌ی ریاضی	توضیحات
MSE (mean squared error)	$\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$	میانگین مجموع مجذور تفاضل مقدار واقعی با مقدار پیش‌بینی شده در مسائل رگرسیون.
MAE (mean absolute error)	$\frac{1}{N} \sum_{i=1}^N y_i - \hat{y}_i $	میانگین مجموع قدرمطلق تفاضل مقدار واقعی با مقدار پیش‌بینی شده در مسائل رگرسیون.

¹ Loss function

² Optimizer algorithms

Binary Cross Entropy	$-\frac{1}{N} \sum_{i=1}^N y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)$	از توابع خطا مرسوم برای مسائل دسته‌بندی دوتایی ^۱ .
Categorical Cross Entropy	$-\sum_{j=1}^k y_j \log \hat{y}_j$	بسط یافته‌ی تابع binary cross entropy است و خطا را برای k کلاس محاسبه می‌کند.

یکی از مشکلات اساسی در آموزش شبکه‌های عصبی، مشکل گم‌شدن گرادینان در شبکه‌های عمیق است. در شبکه‌های عمیق، زمانی که اطلاعات از لایه به لایه برای آموزش منتقل می‌شوند، گرادینان‌ها ممکن است به گونه‌ای کاهش یابند که به سختی یا حتی ناتوانایی در آموزش لایه‌های عمیق‌تر منجر می‌شود. به منظور رفع این مشکل ساختارهای متنوعی معرفی شده‌اند. یکی از این ساختارها، ساختار Residual است. همانطور که در شکل ۷ نشان داده شده است، در ساختار Residual، این مشکل با افزودن یک مسیر مستقیم (یا اتصال مابقی) بین ورودی و خروجی لایه‌ها حل می‌شود. به عبارت دیگر، خروجی لایه جدید با ورودی لایه قبلی جمع می‌شود. این اتصال مستقیم به مدل این امکان را می‌دهد تا اطلاعات را به سرعت از لایه به لایه منتقل کرده و مشکل گم‌شدن گرادینان را کاهش دهد.



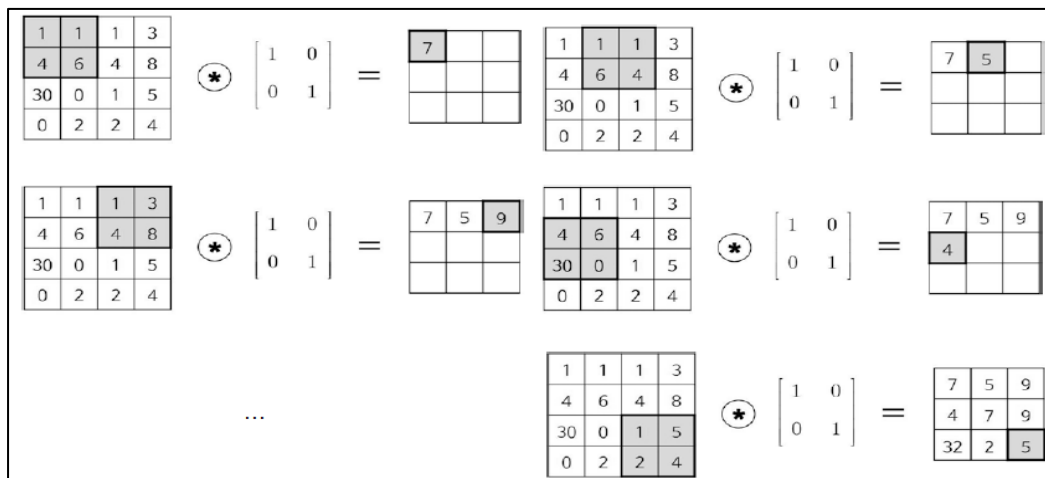
شکل ۷: ساختار residual باعث جلوگیری از گم‌شدن گرادینان می‌کند [21].

در ادامه به یک سری از شبکه‌های عصبی متداول و روش‌های معروف در این حوزه، پرداخته می‌شود.

¹ Binary classification

۲-۳-۲- شبکه‌های عصبی پیچشی^۱

شبکه‌های عصبی پیچشی یا CNNها، نوع دیگری از شبکه‌های عصبی هستند که معمولاً برای استخراج ویژگی‌ها و استخراج الگوها استفاده می‌شود. این شبکه‌ها معمولاً برای پردازش تصاویر و پردازش داده‌های توالی استفاده می‌شوند. شبکه‌های عصبی پیچشی از لایه‌های پیچشی^۲ و لایه‌های ادغام^۳ تشکیل شده‌اند. لایه‌های پیچشی به عنوان فیلترها یا کرنل‌ها عمل کرده و با حرکت این فیلترها روی داده‌ها، ویژگی‌های مختلف با محاسبه‌ی عملیات پیچشی^۴ از آن استخراج می‌کنند. عملیات پیچشی، حاصل جمع ضرب هر مؤلفه کرنل با مؤلفه نظیرش در داده می‌باشد. لایه‌های ادغام به وسیله‌ی روش‌های تجمیع^۵، با کاهش ابعاد داده و حذف اطلاعات غیرضروری، کمک به کاهش تعداد پارامترها و افزایش سرعت آموزش می‌کنند [15] [14]. لایه‌های پیچشی به دو دسته‌ی یک بعدی و دو بعدی تقسیم‌میشوند. کرنل‌های لایه‌ی پیچشی دو بعدی در دو جهت حرکت می‌کنند و بیشتر برای پردازش تصویر کاربرد دارند. شکل ۸، عملیات پیچشی دو بعدی را نشان می‌دهد. در مقابل، کرنل‌های لایه‌ی پیچشی یک بعدی تنها در یک بعد حرکت می‌کنند و بیشتر برای پردازش داده‌های توالی، مثل پردازش سیگنال یا پردازش متن، استفاده می‌شوند. شکل ۹ نحوه‌ی حرکت کرنل پیچشی یک بعدی را نشان می‌دهد. در فرایند آموزش این شبکه‌ها، مؤلفه‌های شبکه بروزسانی شده و الگوهای مناسب از داده‌ها استخراج می‌شود.



شکل ۸: مراحل بدست آوردن خروجی از اعمال عملیات پیچشی یک فیلتر بر روی یک ماتریس ورودی [24].

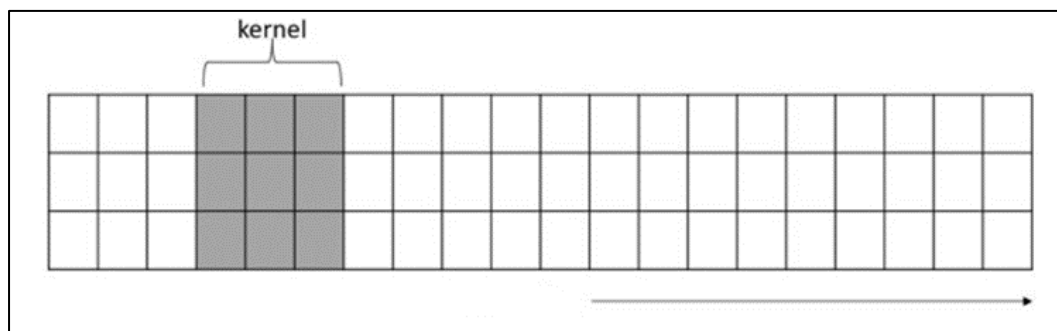
¹ Convolutional neural network

² Convolutional layers

³ Pooling layers

⁴ Convolution

⁵ Aggregation methods



شکل ۹: نحوه‌ی حرکت کرنل پیچشی یک بعدی [25]

۲-۳-۳- شبکه‌های عصبی بازگشتی^۱

شبکه‌های عصبی بازگشتی یا RNNها یک نوع از معماری‌های شبکه‌های عصبی هستند که برای پردازش داده‌های توالی، مانند داده‌های زمانی یا پردازش کلمات در جمله، استفاده می‌شوند. همانطور که در شکل ۱۰ نمایش داده شده است، این نوع از شبکه‌های عصبی از ویژگی حافظه‌ی بازگشتی^۲ برخوردار هستند که این امکان را فراهم می‌سازد تا اطلاعات از گذشته در فرایند آموزش و پیش‌بینی مورد استفاده قرار گیرد. این شبکه‌ها، در مسائل مختلفی مانند ترجمه ماشینی^۳، تشخیص گفتار^۴، پیش‌بینی متن^۵، و حتی تولید متون^۶ جدید مؤثر هستند. محو شدن اطلاعات داده‌های اولیه و مشکل حذف گرادیان از مشکلات جدی این نوع معماری می‌باشد، با این حال با تغییرات و اصلاح طراحی‌های موجود در این شبکه‌ها، بهبودهای چشمگیری در مدیریت این چالش‌ها ایجاد شده است. LSTM^۷ و GRU^۸ که در به نمایش درآمده‌اند، از متداول‌ترین شبکه‌های عصبی بازگشتی می‌باشند که در زمینه‌ی این چالش‌ها بهینه شده‌اند [26]. شکل ۱۱، معماری‌های GRU و LSTM و تفاوت آن‌ها با RNNهای اولیه را نمایش می‌دهد.

¹ Recurrent neural networks

² Recurrent memory

³ Machine translation

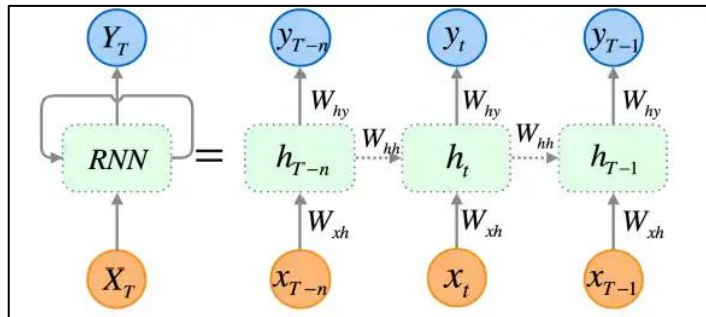
⁴ Speech recognition

⁵ Text prediction

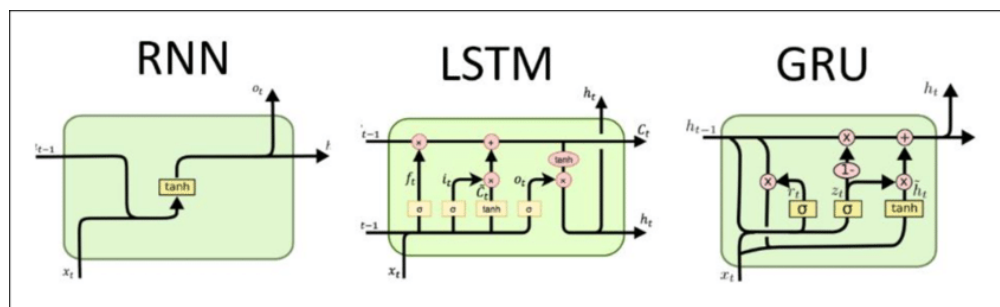
⁶ Text generation

⁷ Long short-term memory

⁸ Gated recurrent unit



شکل ۱۰: نحوه‌ی عملکرد شبکه‌های عصبی بازگشتی [27].



شکل ۱۱: شبکه‌های GRU و LSTM و تفاوت آن‌ها با شبکه‌های RNN اولیه [28].

به منظور مدل‌سازی بهتر داده‌های دنباله‌ای، می‌توان از RNN‌های دوطرفه^۱ (Bi-RNN) استفاده نمود. استفاده از Bi-RNN‌ها می‌تواند در بهبود دقت مدل‌ها در وظایف پردازش زبان طبیعی کمک کند، به خصوص زمانی که وابستگی‌ها و اطلاعات دوطرفه مهم هستند. این شبکه‌ها به خوبی درک ارتباطات دوسویه در داده‌ها و بازایی اطلاعات موثر از هر دو جهت دنباله را فراهم می‌کنند. به عنوان مثال، در تشخیص اسامی یا موجودیت‌ها در یک متن، Bi-RNNs ممکن است نقش مهمی ایفا کنند [29].

۲-۳-۴- ساختار کدگذار-کدگشا^۲

ساختارهای مختلفی برای حل مسائل متفاوت بر بستر یادگیری عمیق ارائه شده‌اند. یکی از این ساختارها، ساختار کدگذار-کدگشا است. ساختار کدگذار-کدگشا یک معماری متداول در حوزه شبکه‌های عصبی است که برای مسائل مختلفی از جمله ترجمه ماشینی، تولید متن، و تولید تصویر^۳ مورد استفاده قرار می‌گیرد. کدگذار مسئول تبدیل ورودی به یک فضای نهان^۴ یا بردار ویژگی است.

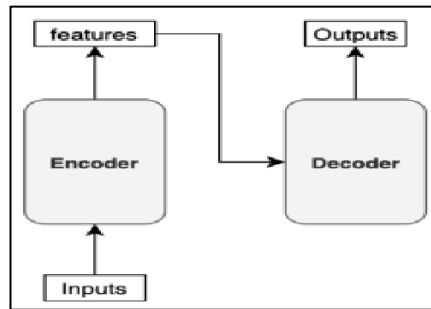
¹ Bidirectional recurrent neural networks

² Encoder-decoder structure

³ Image generation

⁴ Latent space

این فضای نهان معمولاً اطلاعات مهم و تجمعی از ویژگی‌های ورودی را نمایندگی می‌کند. سپس کدگشا با استفاده از این نمایش نهان، وظیفه بازسازی و یا تولید خروجی مورد نظر را برعهده دارد [30]. شکل ۱۲، ساختار کلی کدگذار-کدگشا را نشان می‌دهد.



شکل ۱۲: ساختار کدگذار-کدگشا [31]

۲-۳-۵- روش‌های مبتنی بر توجه ۱ و معماری ترنسفرمر ۲

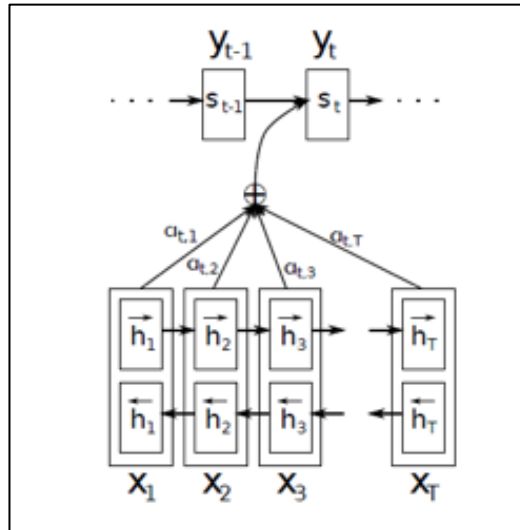
ترنسفرمر یک معماری مدل عمیق یادگیری ماشین است که به عنوان یکی از پیشرفت‌های قابل توجه در زمینه پردازش زبان طبیعی و وظایف مرتبط با آن شناخته می‌شود و معمولاً به عنوان یک انقلاب در مدل‌های زبانی و توانایی پردازش اطلاعات دنباله‌ای شناخته می‌شود. یکی از ویژگی‌های اصلی ترنسفرمر، استفاده از ساختار توجه است که به مدل این امکان را می‌دهد تا به مرور دنباله ورودی خود، توجه متفاوتی به بخش‌های مختلف از داده داشته باشد. این ساختار توجه، توانمندی مدل را در درک و استخراج اطلاعات مهم از داده‌های ورودی افزایش می‌دهد. ترنسفرمر به دلیل قابلیت پردازش دنباله‌های طولانی، امکان توجه به ارتباطات دور و ابعاد بالا، به خصوص در وظایفی که وابستگی‌های طولانی دارند مانند ترجمه ماشینی و تولید خودکار متن، مورد توجه قرار گرفته است [5].

همانطور که گفته شده، ساختار توجه، اصلی‌ترین ویژگی معماری ترنسفرمرها می‌باشد. همه‌ی قسمت‌های یک دنباله، در خروجی به یک اندازه اهمیت ندارند. روش‌های مبتنی بر توجه در مسائلی مثل خلاصه‌سازی متن، فهم متون، دسته‌بندی متون کاربرد فراوانی دارند [32]. مکانیزم توجه یک مفهوم کلیدی در زمینه یادگیری عمیق و پردازش زبان طبیعی است که به شبکه‌های عصبی اجازه می‌دهد تا به نقاط مهم یا ویژگی‌های مشخصی در ورودی خود تمرکز کنند. برای مثال، در پردازش زبان طبیعی، مکانیزم توجه به مدل‌ها امکان تمرکز بیشتر بر روی کلمات مهم یا جملات کلیدی را می‌دهد. این توانایی می‌تواند به بهبود کارایی مدل در ترجمه ماشینی، پرسش و پاسخ و وظایف دیگر

¹ Attention-based methods

² Transformer

کمک کند. این مفهوم ابتدا در سال ۲۰۱۵، در معماری ترنسفرمر معرفی شد و از آن زمان به عنوان یک عنصر اساسی در بسیاری از مدل‌های پیشرفته عصبی مورد استفاده قرار گرفته است [5]. شکل ۱۳ ساختار توجه در فرایند بدست آوردن نمایش بهتر برای دنباله را نشان می‌دهد.



شکل ۱۳: نمایش مکانیزم توجه برای بدست آوردن نمایش برای یک دنباله [33]

در هر سلول توجه^۱، سه مفهوم کلید^۲، مقدار^۳ و پرسش^۴ وجود دارد. در هر لایه با استفاده از یک معیار شباهت، میزان شباهت بین کلید و پرسش محاسبه می‌شود. خروجی هر عنصر متناظر با ورودی به صورت معادله‌های ۱ تا ۳ محاسبه می‌شود [5].

$$a_{ij} = \text{SIM}(q_i, k_j) \quad (1)$$

$$a'_{ij} = \frac{e^{a_{ij}}}{\sum_k a_{ik}} \quad (2)$$

$$\hat{v}_i = a'_{ij} v_j \quad (3)$$

در عبارات ریاضی بالا منظور از q پرسش، k کلید و v مقدار می‌باشد. تابع SIM ، یک تابع شباهت می‌باشد. معادله‌ی ۲ نیز نحوه‌ی محاسبه‌ی SoftMax است که در آن ضریب اهمیت،

¹ Attention cell

² key

³ value

⁴ query

محاسبه می‌شود. یکی از روش‌های محاسبه‌ی شباهت می‌تواند ضرب داخلی مقیاس‌بندی شده^۱ باشد. معادله ۴ نحوه‌ی محاسبه‌ی ضرب داخلی مقیاس‌بندی شده را نشان می‌دهد [5].

$$f_{dot}(Q, K) = \frac{QK^T}{\sqrt{d}} \quad (4)$$

در عبارت ریاضی بالا منظور از Q و K دو بردار برای محاسبه‌ی شباهت هستند. d بعد بردارهای Q و K می‌باشد.

استفاده همزمان و موازی از چند سلول توجه می‌تواند ویژگی‌های بهتر از دیدگاه‌های مختلف تولید کند. به ساختار استفاده چند سلول توجه به صورت همزمان و موازی، توجه چند سره^۲ می‌گویند [5]. همچنین اگر در ساختار توجه، کلید و مقدار یکسان باشند و با استفاده از دنباله‌ی ورودی، سعی در ایجاد نمایش از همان دنباله شود، ساختار خود توجه^۳ گویند [34].

۲-۳-۶- شبکه‌های عصبی گراف^۴

در ریاضیات، گراف را می‌توان به عنوان یک جفت مرتب (V, E) تعریف کرد، که در آن V مجموعه‌ای از گره‌ها^۵ (نقاط) و E مجموعه‌ای از یال‌ها^۶ (ارتباطات) است. هر یال با دو گره مختلف از V مرتبط می‌شود. این تعریف مشخص می‌کند که گراف چگونه اجزا و ارتباطات بین آنها را نمایش می‌دهد [35]. همچنین گراف‌ها ابزارهای قدرتمندی در زمینه هوش مصنوعی هستند که در تحلیل و نمایش روابط بین اشیاء و اطلاعات مورد استفاده قرار می‌گیرند. گراف‌ها به تدریج به عنوان یک مدل انتزاعی جهت نمایش روابط میان داده‌ها و اجزای سیستم‌ها مورد توجه قرار گرفته‌اند. در حوزه هوش مصنوعی، گراف‌ها به عنوان یک ساختار داده‌ای اساسی مورد استفاده قرار می‌گیرند. این ساختارها قابلیت نمایش روابط پیچیده میان داده‌ها را فراهم می‌کنند و در بسیاری از مسئله‌ها مانند معماری داده^۷، شبکه‌های اجتماعی^۸، پردازش زبان طبیعی، مسائل مسیریابی^۹، و تجزیه و تحلیل شبکه‌های پیچیده^{۱۰} کاربرد دارند [36]. همانطور که در شکل ۱۴ نشان داده شده است، گراف می‌تواند

¹ Scaled dot-product

² Multi-head attention

³ Self-attention

⁴ Graph neural networks

⁵ Nodes

⁶ Edges

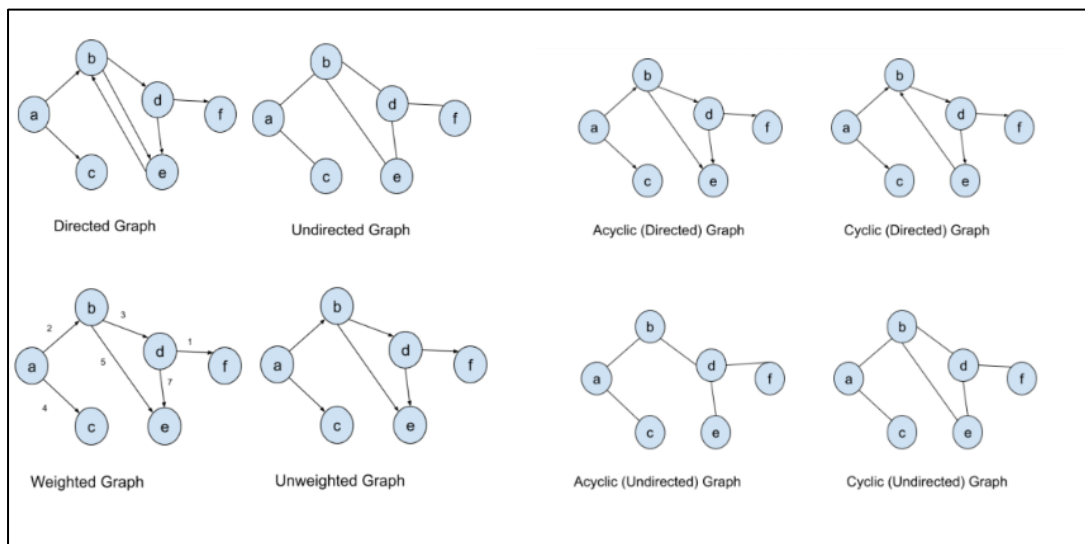
⁷ Data architecture

⁸ Social networks

⁹ Routing problems

¹⁰ Analysis of complex networks

جهت‌دار^۱ یا بی‌جهت^۲، وزن‌دار^۳ یا بی‌وزن^۴، گراف دوره‌دار^۵ و بدون دور^۶ باشد و تبع آن، مفاهیم متنوعی از گراف‌ها ایجاد می‌شود [35]. همچنین اگر گره‌های یک گراف ماهیت یکسان داشته باشند، به آن گراف همگن^۷ و اگر ماهیت گره‌ها یا هم یکسان نباشند به آن گراف ناهمگن^۸ گفته می‌شود.



شکل ۱۴: گراف‌ها می‌توانند وزن‌دار یا بی‌وزن، جهت‌دار یا بی‌جهت، دوردار یا بی‌دور باشند [۲۵]

شبکه‌های عصبی گراف یک دسته از مدل‌های یادگیری عمیق هستند که برای مدل‌سازی داده‌های گرافی مانند شبکه‌های اجتماعی، شبکه‌های مولکولی و ساختارهای گرافی مشابه به کار می‌روند. این شبکه‌ها توانمندی بالایی در تفسیر و تحلیل روابط پیچیده و الگوهای شبکه‌های گرافی دارند. به عبارت دیگر، آنها قابلیت مدل‌سازی تعاملات و وابستگی‌های بین گره‌ها و لینک‌های یک گراف را دارا هستند [37].

شبکه‌های عصبی گراف از مکانیزم انتقال پیام^۹، به منظور تبادل اطلاعات بین گره‌ها در یک گراف بهره می‌برند. در این مکانیزم، هر گره اطلاعات مرتبط با خود را با همسایگان خود به اشتراک می‌گذارد و از اطلاعات جمع‌آوری شده برای به‌روزرسانی ویژگی‌های خود استفاده می‌کند. این فرآیند

¹ Directed graph

² Undirected graph

³ Weighted graph

⁴ Unweighted graph

⁵ Cyclic graph

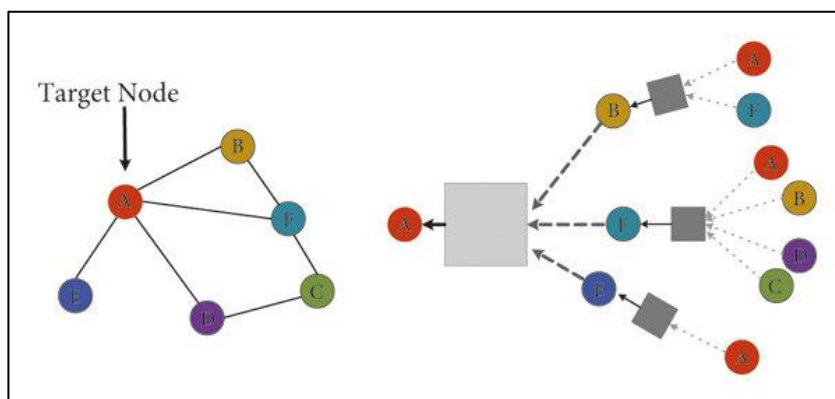
⁶ Acyclic graph

⁷ Homogeneous graph

⁸ Heterogeneous graph

⁹ message passing

چند مرحله‌ای ادامه می‌یابد تا اطلاعات به صورت تدریجی از یک گره به دیگری منتقل شوند. همچنین به غیر از اطلاعات درون یک گره، میتوان برای یال‌ها نیز اطلاعات و ارتباطاتی^۱ نیز در نظر گرفت. مدیریت پیچیدگی مدل‌ها، تعامل با ساختار گراف، انعطاف پذیری در مقابل اندازه‌ی گراف‌ها، آموزش ناهمگن^۲ گراف‌ها و مقاومت در برابر نویز از قابلیت‌های شبکه‌های عصبی می‌باشد [37]. یکی از اولین و متداول‌ترین شبکه‌های عصبی گراف، GCN^۳ها^۴ می‌باشند که در شکل ۱۵ نحوه‌ی عملکرد آن نمایش داده شده است. یکی دیگر از شبکه‌های عصبی گراف معروف، شبکه‌ی عصبی GAT^۴ می‌باشد که با استفاده از مکانیزم توجه، می‌تواند میزان اهمیت بین گره‌ها در فرایند انتقال پیام را آموزش ببیند. همچنین معماری SageConv یک نوع از معماری های شبکه‌های عصبی GraphSAGE^۵ است، که با جمع‌آوری اطلاعات از همسایگان هر گره، نمایش‌های گره‌ها را یاد می‌گیرد و به عنوان جانشانی گره‌ها^۶ برای مدل‌سازی اطلاعات پیکربندی^۷، از آن استفاده می‌شود. این معماری از یک عامل پیچشی استفاده می‌کند که به آن این امکان را می‌دهد که ویژگی‌های پیچیده‌تری را درک کند. این مدل به افراد اجازه می‌دهد نمایش‌های برداری برای گره‌های جدید و دیده نشده در گراف تولید کند، به جای آن که به طور مستقیم بر روی گره‌های موجود در مرحله آموزش آن یادگیری شود [38].



شکل ۱۵: نحوه‌ی انتقال پیام برای یک گره مشخص در شبکه‌ی GCN [39]

¹ relations

² Heterogeneous Learning

³ Convolutional network

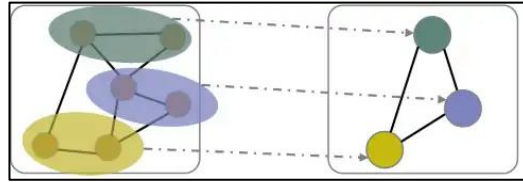
⁴ Graph attention networks

⁵ Graph sample and aggregated embeddings

⁶ Node embedding

⁷ topology

برخی مواقع در شبکه‌های عصبی گراف از لایه‌های ادغامی گراف^۱ به منظور به‌دست آوردن گراف کوچک‌تر استفاده می‌شود. همچنین با این روش می‌توان نمایشی برای گراف مورد نظر نیز به‌دست آورد. در شکل ۱۶ یک نمونه از ادغام گره‌های گراف نمایش داده شده است.



شکل ۱۶: ادغام گره‌های گراف و تولید گراف کوچک‌تر با استفاده از لایه‌ی ادغام گراف [40]

۲-۴- جانشانی کلمه^۲

در پردازش زبان طبیعی، برای استفاده از روش‌های یادگیری ماشین، کلمات یا عبارات باید به بردارهای عددی تبدیل شوند. بنابراین روشی برای تبدیل متن به بردار نیاز است. جانشانی کلمه یک روش در زمینه پردازش زبان طبیعی است که اهمیت زبانی کلمات را در یک فضای چندبعدی عددی نمایش می‌دهد. این تکنیک مفهومی از معانی و ارتباطات بین واژگان را در یک فضای چندبعدی مدل می‌کند. به این ترتیب، واژگان مشابه در این فضا به نزدیکی هم نقل می‌کنند و این اطلاعات معنایی برای الگوریتم‌های پردازش زبان طبیعی بسیار حائز اهمیت است. برای مثال، اگر دو کلمه معنایی مشابهی داشته باشند، مطابق با این روش، بردارهای متناظر با این کلمات نیز به هم نزدیک خواهند بود. این امر به الگوریتم‌ها کمک می‌کند تا معانی و ارتباطات میان واژگان را درک و استخراج کنند [41]. از روش‌های معروف در این حوزه می‌توان به word2vec، glove^۳ و برخی مدل‌های از پیش آموزش دیده شده^۴ مثل مدل BERT^۵ نام برد.

۲-۴-۱- جانشانی کلمه Glove

جانشانی کلمه Glove یک روش برای ایجاد بردارهای کلمات در پردازش زبان طبیعی است. این روش توسعه داده شده توسط پژوهشگران گروه گوگل به منظور بهبود نمایش واژگان در فضای چندبعدی می‌باشد. در این روش، بردارهای کلمات بر اساس آمارهای مشاهده شده در متون آموزش

¹ Graph pooling layers

² Word embedding

³ Global vectors for word representation

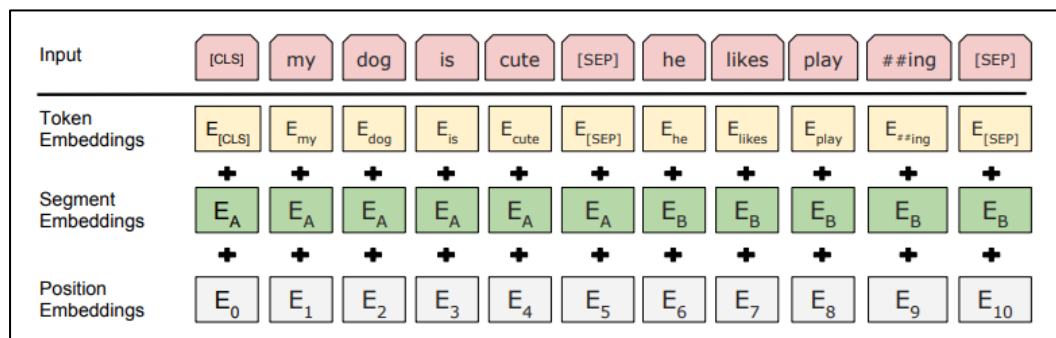
⁴ Pretrained model

⁵ Bidirectional encoder representations from transformers

دیده می‌شوند. الگوریتم Glove به این صورت عمل می‌کند که بردارهای کلمات را طوری بهینه می‌سازد که معانی و ارتباطات معنایی بین واژگان در فضای چندبعدی به خوبی بازنمایی شوند. این روش تاکید بر توزیع احتمالاتی کلمات در متون و سعی در بازنمایی معانی ضمنی و روابط معنایی میان واژگان را دارد [42]. Glove به عنوان یکی از مدل‌های موفق در زمینه جانمایی کلمه به شناخته شده است و به طور گسترده در پروژه‌های پردازش زبان طبیعی مورد استفاده قرار گرفته است.

۲-۴-۲- مدل BERT

مدل BERT یکی از مدل‌های برجسته در زمینه پردازش زبان طبیعی است که توسط گوگل منتشر شده است. این مدل بر اساس معماری ترنسفورمر ساخته شده و به دلیل توانایی بازنمایی بردارهای کلمات در متن به صورت دوطرفه^۱، از دقت واژگان بالایی برخوردار است. BERT قابلیت انجام ماموریت‌های مختلف در زمینه پردازش زبان طبیعی، توانایی درک ارتباطات و مفاهیم پیچیده در یک متن را داراست. برای استفاده از BERT، متن ورودی به شکل تکه‌های کوچکتر^۲ تقسیم می‌شود و سپس با استفاده از لایه‌های ترنسفورمر، بردارهای بازنمایی متن به دست می‌آیند. این بردارها سپس می‌توانند برای ماموریت‌های متنوعی از جمله تشخیص ماهیت‌های نامدار^۳، استخراج ارتباطات^۴، ترجمه ماشینی، و حتی فهم متن وابسته به سوالات^۵ استفاده شوند. همانطور که در شکل ۱۷ آمده است، این مدل از سه جانمایی توکن، بخش و موقعیت، برای بدست آوردن جانمایی کلمه استفاده می‌کند. این معماری و آموزش آن بر روی مقدار زیادی داده باعث شده است تا جانمایی کلمات با استفاده از ویژگی‌های زمینه‌ای^۶ استخراج شوند [43].



شکل ۱۷: جانمایی کلمه در مدل BERT با استفاده از سه جانمایی توکی، بخش و موقعیت تولید می‌شود.

¹ bidirectional

² token

³ Named entity recognition (NER)

⁴ Relation extraction (RE)

⁵ Question answering (QA)

⁶ Contextual feature

۲-۵- معیارهای شباهت بین بردارها

در یادگیری ماشین، معیارهای شباهت بین بردارها از اهمیت چشمگیری برخوردارند. این معیارها به تعیین میزان تشابه یا فاصله بین داده‌ها کمک می‌کنند و در مسائل مختلفی از جمله دسته‌بندی، خوشه‌بندی، یادگیری نظارت‌نشده، و بازیابی اطلاعات مورد استفاده قرار می‌گیرند. برای اندازه‌گیری شباهت بین دو بردار، معیارهایی مانند فاصله اقلیدسی^۱ (معادله ۵) یا شباهت کسینوسی^۲ (معادله ۶) استفاده می‌شود. این معیارها می‌توانند به مدل‌های یادگیری ماشین کمک کنند تا الگوها و ارتباطات پنهان در داده‌ها را درک کرده و اطلاعات مفیدی را استخراج نمایند.

استفاده از معیارهای شباهت در یادگیری ماشین به ما این امکان را می‌دهد تا الگوها و ارتباطات پنهان در داده‌ها را بهبود بخشیم و به دقت و کارایی مدل‌ها افزوده و اطلاعات بیشتری از داده‌ها استخراج کنیم.

$$\text{euclidean}_{v_i, v_j} = \sqrt{\sum_{k=1}^N |v_{ik} - v_{jk}|^2} \quad (5)$$

$$\cos_{v_i, v_j} = \frac{\sum_{k=1}^N v_{ik} v_{jk}}{\sqrt{\sum_{k=1}^N v_{ik}^2} \sqrt{\sum_{k=1}^N v_{jk}^2}} = \frac{v_i \cdot v_j}{|v_i| |v_j|} \quad (6)$$

در محاسبه‌ی شباهت کسینوسی از تقسیم ضرب داخلی^۳ دو بردار به روی اندازه‌ی هر کدام استفاده می‌شود.

۲-۶- شاخص‌های ارزیابی در مسئله‌ی خلاصه‌سازی متن

یکی از چالش‌های موجود در سیستم‌های خلاصه‌سازی متن، چگونگی تعیین میزان کیفیت یک خلاصه‌ی تولید شده توسط سیستم می‌باشد. به منظور مقایسه‌ی میزان کیفیت و درستی خلاصه‌های تولید شده توسط این سیستم‌ها در هر دو رویکرد انتزاعی و استخراجی، پژوهش‌هایی برای پیدا کردن نمایش عددی، انجام شده است. معیارهای ارزیابی یک سیستم خلاصه‌سازی خودکار متن به دو دسته‌ی ذاتی^۴ و بیرونی^۵ تقسیم می‌شوند. معیارهای بیرونی میزان پیروی از یک ساختار (بیشتر مواردی مثل مسائل پاسخ‌دهی به سؤالات^۶ یا بازیابی اطلاعات^۱) را نمایش می‌دهند. اما معیارهای

¹ Euclidean distance

² Cosine similaity

³ Dot product

⁴ Intrinsic

⁵ Extrinsic

⁶ Question answering

ذاتی بر روی انسجام و میزان کیفیت معنایی، تمرکز دارند. معیارهای ذاتی نیز خود به دو دسته‌ی، معیارهای ارزیابی کیفیت متن^۲ و ارزیابی محتوا^۳ تقسیم می‌شوند. در سیستم‌های خلاصه‌سازی متن، معمولاً از معیارهای ذاتی استفاده می‌شود [45], [44]. در ادامه به برخی از معروف‌ترین و پرکاربردترین معیارهای ارزیابی در این سیستم‌ها اشاره خواهد شد.

۲-۶-۱- معیار صحت یا Precision

معیار صحت یا precision تعیین می‌کند که خلاصه‌ی به‌دست آمده، به چه میزان مختصر و فاقد کلمات اضافی می‌باشد. میزان این معیار مطابق با رابطه‌ی ۷ بدست می‌آید [45].

$$precision = \frac{S_{ref} \cap S_{cand}}{S_{cand}} \quad (7)$$

در این رابطه، نسبت تعداد کلمات مشترک خلاصه‌ی مرجع (S_{ref}) و خلاصه‌ی حاصل از سیستم خلاصه‌سازی (S_{cand}) به تعداد کل کلمات خلاصه‌ی بدست آمده، محاسبه می‌شود.

۲-۶-۲- معیار Recall

معیار Recall میزان پوشش خلاصه‌ی مرجع توسط خلاصه‌ی حاصل شده را تعیین می‌کند. مقدار این معیار با استفاده از رابطه‌ی ۸ بدست می‌آید [45].

$$Recall = \frac{S_{ref} \cap S_{cand}}{S_{ref}} \quad (8)$$

۲-۶-۳- معیار F-measure

به منظور ادغام دو معیار Recall و precision، از معیار F-measure استفاده می‌شود. رابطه‌ی ۹ مقدار این معیار را به دست می‌آورد [45].

$$F_{measure} = \frac{2(Precision)(Recall)}{Precision + Recall} \quad (9)$$

۲-۶-۳- معیار ROUGE^۴

معیار ROUGE مرسوم‌ترین معیار ارزیابی در سیستم‌های خلاصه‌سازی خودکار متن است. معیار ROUGE با استفاده از چندتایی‌های واژگان متوالی (N-grams) مشترک بین خلاصه‌ی مرجع و

¹ Information retrieval

² Text quality evaluation

³ Content evaluation

⁴ Recall-Oriented Understudy for Gisting Evaluation

خلاصه‌ی تولید شده، یک مقدار عددی برای کیفیت محتوای خلاصه ارائه می‌دهد. رابطه‌ی ۱۰ میزان عددی ROUGE متناسب با تعداد واژگان متوالی موردنظر را بدست می‌آید [46].

$$Rouge(n)_{Recall} = \frac{count((\forall gram_n \in source) \cap (\forall gram_n \in reference))}{count(\forall gram_n \in reference)} \quad (10)$$

رابطه‌ی فوق مقدار Rouge با رویکرد Recall را نشان می‌دهد که نسبت مجموع n-gram‌های مشترک در جملات خلاصه‌ی مرجع و خلاصه‌ی تولید شده ($count(gram_n)$) به مجموع n-gram مشترک در جملات خلاصه‌ی مرجع ($count(gram_n)$) را بدست می‌آورد [45]. همانطور که در رابطه‌ی ۱۱ آمده است، Rouge با رویکرد Precision محاسبه می‌شود و با توجه به دو معیار Precision و Recall می‌توان Rouge با معیار F-measure را به صورت معادله‌ی ۱۲ محاسبه کرد.

$$Rouge(n)_{Precision} = \frac{count((\forall gram_n \in source) \cap (\forall gram_n \in reference))}{count(\forall gram_n \in source)} \quad (11)$$

$$Rouge(n)_F = \frac{2(Rouge(n)_{Precision})(Rouge(n)_{Recall})}{Rouge(n)_{Precision} + Rouge(n)_{Recall}} \quad (12)$$

در مسئله‌ی خلاصه‌سازی خودکار متن ۳ معیار R1، R2 و RL مرسوم‌ترین شاخص‌های گزارش شده در مقالات می‌باشند. R1 به معنای محاسبه‌ی ROUGE با n برابر ۱ می‌باشد. این شاخص میزان همپوشانی توکن‌های یکتایی^۱ بین خلاصه‌ی سیستم و خلاصه‌ی مرجع را نشان می‌دهد. همچنین R2 به معنای ROUGE با n برابر ۲ است که نشان‌دهنده‌ی همپوشانی توکن‌های دوتایی^۲ بین خلاصه‌ی سیستم و خلاصه‌ی مرجع می‌باشد. RL نیز میزان ROUGE بر اساس طول بزرگترین زیردنباله‌ی مشترک بین دو خلاصه را نشان می‌دهد [45]. برای مثال دو متن "گربه‌ی سیاه روی پنجره نشست." و "کنار پنجره، گربه سیاه نشست." در این مثال کلمات "گربه"، "سیاه"، "پنجره" و "نشست" کلمات مشترک این دو متن می‌باشند. در محاسبه‌ی R1 این چهار کلمه در نظر گرفته می‌شوند. در حالی که تنها عبارت "گربه‌ی سیاه" به عنوان عبارت دوتایی مشترک هستند. در RL نیز توکن‌های یکتا با حفظ ترتیب در نظر گرفته می‌شوند. در این مثال کلمات "گربه"، "سیاه" و "نشست" بزرگ‌ترین زیر دنباله با حفظ ترتیب در دو متن هستند، و برای RL تنها این ۳ کلمه در نظر گرفته می‌شوند.

¹ Unigram

² Bigram

۷-۲- جمع‌بندی

در این فصل به بررسی مفاهیم مورد استفاده در این پژوهش پرداخته شد. ابتدا یادگیری ماشین و یادگیری عمیق و برخی تکنیک‌های آن‌ها معرفی گردید. یکی از بزرگ‌ترین مزیت‌های یادگیری عمیق استخراج ویژگی‌های نهان داده‌های ورودی می‌باشد. شبکه‌های عصبی معمولاً شامل یک لایه‌ی ورودی، لایه‌ی نهان و لایه‌ی خروجی هستند. با استفاده از برخی لایه‌های شبکه‌های عصبی مثل لایه‌های CNN و RNN می‌توان ویژگی‌ها و الگوهای داده‌های توالی را استخراج کرد. همچنین برخی مسائل را می‌توان با استفاده از گراف مدل کرد و شبکه‌های عصبی گراف باعث می‌شوند تا اطلاعات بین گره‌های مجاور با سیاست‌های مختلف، به اشتراک گذاشته و در حل مسائل استفاده شوند. همچنین یکی از ساختارهای معروف شبکه‌های عصبی عمیق ساختار کدگذار-کدگشا است که یک نمایش میانی توسط کدگذار از دنباله ورودی ایجاد می‌شود و کدگشا این نمایش را به دنباله خروجی تبدیل می‌کند. در ادامه به مکانیزم توجه پرداخته شد. مکانیزم توجه یک مفهوم کلیدی در زمینه یادگیری عمیق و پردازش زبان طبیعی است که به شبکه‌های عصبی اجازه می‌دهد تا به نقاط مهم یا ویژگی‌های مشخصی در ورودی خود تمرکز کنند. همچنین به منظور تبدیل کلمات به بردارهای عددی از مدل‌های جانشانی کلمات استفاده می‌شود. از روش‌های معروف در این حوزه می‌توان به glove، word2vec و برخی مدل‌های از پیش آموزش دیده شده مثل مدل BERT نام برد. در آخر نیز، شاخص‌های ارزیابی در مسئله‌ی خلاصه‌سازی خودکار متن معرفی گردید. شاخص ROUGE مرسوم‌ترین شاخص ارزیابی این حوزه می‌باشد. ROUGE انواعی دارد که در بین آن‌ها R1، R2 و RL مهم‌ترین آن‌ها می‌باشد و معمولاً در پژوهش‌های مربوط به حوزه‌ی خلاصه‌سازی متن گزارش می‌شوند. در فصل آینده به بررسی رویکردها و پژوهش‌های انجام شده در خلاصه‌سازی خودکار متن پرداخته شده است.

فصل سوم

پیشینه پژوهش

۳-۱- مقدمه

رویکرد استخراجی خلاصه‌سازی متن را می‌توان به صورت یک مسئله‌ی دسته‌بندی جملات^۱ در نظر گرفت، بدین صورت که باید مدل تشخیص دهد، هر جمله در دسته‌ی خلاصه قرار بگیرد یا خیر [47]. استخراج برچسب‌ها از خلاصه‌ی انتزاعی، با روش برچسب‌گذاری oracle که یک روش مبتنی بر امید ریاضی^۲ است، انجام می‌شود [48]. پس هر مدل استخراجی، از یک دسته‌بند^۳ برای عملیات دسته‌بندی جملات، استفاده می‌کند. این دسته‌بندها می‌تواند به صورت یک شبکه‌ی عصبی (MLP، LSTM و ...) یا به صورت الگوریتم‌هایی مثل الگوریتم‌های رتبه‌بندی^۴ باشد. برای مثال یک الگوریتم رتبه‌بندی می‌تواند از ویژگی‌های آماری، نظیر طول جملات، موقعیت جملات، ویژگی‌های TF-IDF، شباهت‌های کسینوسی و مواردی دیگر، به جملات امتیاز داده و آن‌ها را مرتب سازد [49]. پیش از شبکه‌های عصبی عمیق روش‌هایی با رویکرد یادگیری ماشین یا الگوریتم‌های مبتنی بر منطق فازی وجود داشته است. با رواج پیدا کردن شبکه‌های عصبی، این رویکردها توانست بهبودهای بسیاری در زمینه‌ی خلاصه‌سازی متن بدهد. مدل SummRunNER یکی از اولین مدل‌های مبتنی بر

¹ Sentence classification

² Expectation based method

³ classifier

⁴ Ranking algorithms

شبکه‌های عصبی در زمینه‌ی خلاصه‌سازی بوده است، که توانسته است پیشرفت قابل توجهی بدست آورد [50], [47]. در سال‌های گذشته، مدل‌های مبتنی بر شبکه‌عصبی زیادی در زمینه‌ی خلاصه‌سازی ارائه شده است. با پیدایش شبکه‌ی عصبی BERT و نمایش مبتنی بر زمینه‌ی متن^۱، بهبودهای بسیاری در این زمینه رخ داد [47]. یکی از اصلی‌ترین چالش‌های موجود در این حوزه، در نظر گرفتن روابط موجودیت‌های با فاصله‌ی زیاد است. در این امر، شبکه‌های بازگشتی نتوانسته‌اند موفق ظاهر شوند. به همین منظور، شبکه‌های عصبی گراف و استفاده از گراف در زمینه‌ی خلاصه‌سازی، توانسته‌اند روابط مختلف بین کلمات و جملات و عناصر مختلف متن را در خلاصه‌سازی دخیل کنند. ساخت انواع گراف و تعریف انواع روابط در یک متن به صورت گراف، از دلایل موفقیت شبکه‌های عصبی گراف در این مسائل بوده است [52], [51]. شبکه‌های عصبی گراف به چهار دسته‌ی شبکه‌های عصبی گراف بازگشتی^۲، شبکه‌های پیچشی گراف^۳، خود رمزگذار گراف^۴ و شبکه‌های عصبی گراف فضایی-زمانی^۵ تقسیم می‌شوند [53]. در گراف، گره‌ها به طور طبیعی توسط همسایگان و اتصالات خود تعریف می‌شوند. وظیفه‌ی شبکه‌ی عصبی گراف این است که با مشاهده‌ی اطلاعات موجود در گره‌های مجاور، بردار ویژگی هر گره را به دست آورد [54].

روش‌های موجود در خلاصه‌سازی استخراجی متن به دو دسته‌ی کلی تقسیم می‌شوند:

۱- رویکردهای سنتی ۲- رویکردهای مبتنی بر ماشین لرنینگ.

رویکردهای سنتی در خلاصه‌سازی استخراجی بر اساس قوانین و الگوریتم‌های ثابت عمل می‌کنند. این روش‌ها به استفاده از معیارهایی مانند فراوانی واژگان، ساختار متن، و خصوصیات دیگر متن برای انتخاب جملات مهم متن می‌پردازند. به عنوان مثال، روش‌هایی مبتنی بر محاسبه TF-IDF از تکرار واژگان و اهمیت آن‌ها در متن استفاده می‌کنند.

روش‌های مبتنی بر ماشین لرنینگ از مدل‌های آموزش دیده بر داده‌های آموزشی برای تشخیص جملات مهم و استخراج خلاصه استفاده می‌کنند. این مدل‌ها می‌توانند از شبکه‌های عصبی بازگشتی (RNN)، شبکه‌های عصبی ترنسفرمری (Transformer)، یا روش‌های دیگر بر اساس ماشین لرنینگ استفاده کنند. به عنوان مثال، مدل‌های مبتنی بر ترنسفر مانند BERT می‌توانند با در نظر گرفتن روابط بلندمدت و اطلاعات متن در تصمیم‌گیری برای استخراج جملات کلیدی بهبود بخشند.

¹ Contextual representations

² Recurrent graph neural networks

³ Convolutional graph neural networks

⁴ Graph AutoEncoder

⁵ spatial-temporal graph neural networks

بر اساس ادعاهای موجود در مطالعات، مدل‌های مبتنی بر ماشین لرنینگ اغلب دقت و کارایی بهتری در خلاصه‌سازی استخراجی نسبت به رویکردهای سنتی^۱ ارائه کرده‌اند [50]. این مدل‌ها می‌توانند با توجه به مجموعه‌های داده بزرگ و گسترده، الگوهای پیچیده‌تری را در متون شناسایی کنند و در نتیجه خلاصه‌های دقیق‌تری ایجاد کنند.

در این فصل، به بررسی و مقایسه‌ی معماری مدل‌های استخراجی معروف برپایه‌ی شبکه‌های عصبی در سال‌های اخیر پرداخته می‌شود. در فصل‌های بعد این مدل‌ها به همراه مدل پیشنهادی، با هم مقایسه خواهند شد.

¹ classical approaches

۳-۲- مدل‌های مبتنی بر روش‌های سنتی

در حوزه خلاصه‌سازی استخراجی متن، روش‌های سنتی به استفاده از الگوریتم‌ها و قوانین قابل تعریف برای انتخاب جملات کلیدی از متون بلند می‌پردازند. این روش‌ها بر اساس معیارهایی مانند فراوانی واژگان، ساختار متن، یا خصوصیات دیگر متن، جملات مهم را شناسایی و به عنوان جزوه‌ای از خلاصه نهایی انتخاب می‌کنند.

در بیشتر روش‌های سنتی از روش‌های آماری برای خلاصه‌سازی متن استفاده می‌کنند. در روش‌های آماری ابتدایی برای انتخاب محتوا، اغلب از محاسبه فراوانی کلمات در متن برای شناسایی کلمات مهم مربوط به سند استفاده می‌شود. مکانیزم اصلی این روش‌ها اندازه‌گیری فراوانی هر کلمه در یک سند است و فراوانی کلمات با شواهدی اضافی از جمله فراوانی معکوس کلمه در مجموعه کلی اسناد (مثل ویژگی‌های TF-IDF¹) تنظیم می‌شود. این اقدام به افزایش امتیاز کلمات کم‌فراوان در یک مجموعه کلی و تخصیص امتیازهای کمتر به کلمات خیلی فراوان در متن کمک می‌کند. این روش‌ها فرض می‌کنند که اهمیت یک مفهوم خاص در یک متن به تعداد بارهایی که مفهوم در سند ذکر شده است، درصد متن بستگی دارد. این در حالی است که هر کلمه متفاوت به یک مفهوم مختلف نظیر می‌شود. با این حال، شمارش حضور مفاهیم در متن به دلیل وجود مترادف‌ها (مثل "سگ" و "کودک سگ") و بیان‌های مشارکتی (مثل "اوباما" و "رئیس‌جمهور") که به اتصال متون کمک می‌کنند، کار ساده‌ای نیست. بعد از شناسایی کلمات کلیدی در سند، جملات حاوی این کلمات می‌توانند با استفاده از روش‌های مختلف امتیازدهی و رتبه‌بندی جملات (مانند ارتباط جمله با تعداد کلمات کلیدی آن) انتخاب شوند. همچنین جایگاه جملات در متن نیز به عنوان نشانه‌ای از اهمیت جمله در نظر گرفته می‌شود. به عنوان مثال، در داستان‌های خبری، پاراگراف اول یا ابتدایی معمولاً شامل اطلاعات اصلی در مورد رویداد گزارش شده در خبر است، در حالی که بخش‌های دیگر متن جزئیات و همچنین اطلاعات زمینه‌ای در مورد رویداد را ارائه می‌دهد. بنابراین، انتخاب جملات از ابتدای متن می‌تواند یک استراتژی مناسب باشد. با این حال، در متون علمی معمولاً مقدمه اطلاعات زمینه‌ای را ارائه می‌دهد، در حالی که توسعه‌های اصلی در نتیجه‌گیری‌ها گزارش می‌شود، بنابراین استراتژی مکانی باید به نوع متنی که قرار است خلاصه شود، سازگار شود [55]. اهمیت جایگاه جملات در متون رسمی به قدری است که مدلی با نام Lead-3 که یک مدل ساده و تنها انتخاب سه جمله‌ی اول به عنوان خلاصه است، توانسته است نتایج خوبی به دست آورد [56]. یکی دیگر از روش‌های ساده در رویکردهای سنتی، بررسی اهمیت جملات از طریق مقایسه آن‌ها با عنوان سند خلاصه شده است. به عنوان مثال، یک جمله که حاوی کلمات عنوان باشد، به عنوان مرتبط در نظر گرفته می‌شود و هر چه

¹ Term frequency-inverse document frequency

کلمات عنوان بیشتر باشد، جمله به میزان بیشتری مرتبط خواهد بود [55]. با وجود این که روش‌های جدید با رویکردهای یادگیری ماشین و یادگیری عمیق، پیشرفت‌های زیادی در زمینه‌ی خلاصه‌سازی متن بدست آورده‌اند، ولی همچنان روش‌های سنتی نیز مورد استقبال هستند و از آن‌ها در روش‌های جدیدتر نیز استفاده می‌شوند.

۳-۳- مدل‌های مبتنی بر یادگیری ماشین

با ظهور روش‌های یادگیری ماشین، پیشرفت در زمینه‌های مختلف از جمله رگرسیون، خوشه‌بندی، و دسته‌بندی به شدت افزایش یافت. همانطور که پیش‌تر بیان شد، خلاصه‌سازی متن را می‌توان به عنوان یک مسئله دسته‌بندی تعریف کرد. با توجه به حجم بالای داده‌های موجود امروزی و روش‌های موثر جمع‌آوری مجموعه‌دادگان برای خلاصه‌سازی خودکار متن، روش‌های یادگیری ماشین با موفقیت‌های چشمگیری در این زمینه مواجه شده‌اند [57].

کانداسامی و همکاران از تکنیک یادگیری ماشین برای دسته‌بندی هرزنانه^۱ در توییتر استفاده کردند. در این پژوهش از SVM و الگوریتم Naive Bayes به منظور دسته‌بندی جملات به دو دسته‌ی هرزنانه و غیر هرزنانه استفاده شده است [58]. سیلوا و همکاران نیز، آزمایش‌هایی با روش Naive Bayes روی مجموعه داده CNN-Corpus انجام داده‌اند و در خلاصه‌های استخراجی متن موفق ظاهر شدند [59].

استفاده از روش KNN نیز به عنوان دسته‌بند در این گونه مسائل جایگاه خود را تثبیت کرده است. برای مثال در مقاله "الگوریتم KNN با استفاده از شباهت بین ویژگی‌ها برای خلاصه‌سازی متن" یک نسخه خاص از الگوریتم KNN ارائه می‌دهد که در آن شباهت بین بردارهای ویژگی و موقعیت مکانی جملات محاسبه می‌شود [60].

فتاح و همکاران نیز یک رویکرد جدید یادگیری ماشین برای بهبود انتخاب محتوا در خلاصه‌ساز خودکار متون چند سنده ارائه داده است. این روش از یک مدل خلاصه‌ساز آموزش‌پذیر استفاده می‌کند که ویژگی‌های متنوعی از جملات و پاراگراف‌ها را در نظر می‌گیرد، از جمله شباهت کلمات، ساختار متن، عبارات کلیدی، و اطلاعات غیرضروری. این ویژگی‌ها ترکیب شده و با استفاده از مدل‌های آماری مثل مدل آنتروپی حداکثر^۲، دسته‌بند Naive Bayes و ماشین بردار پشتیبان، مدل‌های خلاصه‌سازی ساخته می‌شوند. نهایتاً، این مدل‌ها به یک مدل ترکیبی تبدیل شده و جملات را بر اساس اهمیت رتبه‌بندی می‌کند [61].

¹ Spam detection

² Maximum entropy model

همچنین با به وجود آمدن روش‌های جدید، جانشانی کلمات قوی‌تر و ترکیب آن‌ها با روش‌های یادگیری ماشین ابتدایی، توانست با پیچیدگی محاسباتی کمتر نسبت به آموزش شبکه‌های عصبی، نتایج خوبی را بدست آورد. در سال ۲۰۱۹، میلر و همکاران، توانستند با ترکیب مدل BERT به عنوان جانشانی کلمه و الگوریتم k-means به عنوان مدل خوشه‌بندی، یک نرم‌افزار به منظور خلاصه‌سازی سخنرانی‌ها تولید کنند و در صنعت به موفقیت‌های چشمگیری برسند [62].

همچنین می‌توان مسئله‌ی خلاصه‌سازی متن را به صورت یک مسئله‌ی بندیت^۱ در حوزه‌ی یادگیری تقویتی در نظر گرفت. در این مسئله یک عامل باید تصمیم‌های توزیع‌شده را انتخاب کند تا بهترین پاداش را برای عملکرد خود بدست آورد. این مسأله به یک شکل ساده‌تر از مسائل یادگیری تقویتی مطرح و به عنوان یک مسأله تصمیم‌گیری توزیعی در نظر گرفته می‌شود. در یک مسئله بندیت، عامل با مواجهه با مجموعه‌ای از گزینه‌ها یا "باندها" که هرکدام دارای یک توزیع احتمال مشخصی از پاداش‌ها هستند، باید تصمیم بگیرد که کدام باند را انتخاب کند. هدف این عامل بهینه کردن جمعیت پاداش کلی از طریق انتخاب بهترین باندها با استفاده از تجربیات خود است. در خلاصه‌سازی با این رویکرد، هر سند یک زمینه است و هر زیرمجموعه مرتب‌شده از جملات یک سند یک عمل متفاوت است. در این تحقیق، یک روش نوآورانه برای آموزش شبکه‌های عصبی به منظور انجام خلاصه‌سازی استخراجی روی اسناد تک‌سنده بدون نیاز به برچسب‌های استخراجی تولید شده به شیوه فوق ابتکاری^۲ معرفی می‌شود. در این رویکرد استخراج خلاصه را به عنوان یک مسئله بندیت در نظر می‌گیرد. در این سناریو، مدل یک سند را به عنوان ورودی جهت خلاصه‌سازی دریافت می‌کند و باید یک دنباله از جملات را برای درج در خلاصه انتخاب کند. الگوریتم تقویت یادگیری با رویکرد مسأله بندیت برای آموزش مدل به منظور انتخاب دنباله‌های جملاتی که امتیاز ROUGE را بیشینه می‌کنند، به کار گرفته می‌شود [63].

۳-۴- مدل‌های مبتنی بر یادگیری عمیق

تا قبل از ظهور یادگیری عمیق، روش‌های سنتی بر مبنای مدل‌های زبانی و ویژگی‌های انتخاب شده، به طور محدود به پیشرفت‌های محدودی دست یافتند. با این حال، این روش‌ها همچنان مشکلاتی از قبیل توانایی ضعیف در درک معانی عمیق متون و تولید خلاصه با پوشش بهتر متن، داشتند.

¹ Bandit Problem

² Heuristic

در دهه‌های اخیر، با پیشرفت مداوم نظریه و فناوری شبکه‌های عصبی، یادگیری عمیق به عنوان یکی از روش‌های موثر در زمینه‌ی خلاصه‌سازی متن شناخته شده است. تقسیم‌بندی‌های متنوعی برای مدل‌های مبتنی بر یادگیری عمیق می‌توان در نظر گرفت. در ادامه به بررسی چند رویکرد و ایده‌های اصلی در این حوزه پرداخته می‌شود.

۳-۴-۱ - مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی

در حوزه خلاصه‌سازی متن، مدل‌های استخراجی مبتنی بر شبکه‌های عصبی بازگشتی (RNN) ابزارهای موثری برای خلاصه‌سازی خودکار و تولید خروجی هستند. این مدل‌ها عمدتاً از معماری‌هایی مانند شبکه‌های LSTM و یا GRU بهره می‌برند. در طراحی این مدل‌ها، معمولاً یک شبکه بازگشتی به عنوان کدگذار برای تجسم ویژگی‌های مهم متون و یک مکانیزم انتخاب برای تمرکز بر نقاط کلیدی متن به کار می‌روند. برای مثال، یک مدل RNN می‌تواند با دریافت ویژگی‌های هر جمله به عنوان ورودی، توانایی انتخاب جملات مهم و تولید یک خلاصه فشرده از متن را داشته باشد. همچنین، شبکه‌های RNN دوطرفه نیز مورد استفاده قرار می‌گیرند تا به مدل اطلاعات کامل‌تری از متن دسترسی داده شود [64], [65].

یکی از اولین مدل‌های مبتنی بر شبکه عصبی در زمینه‌ی خلاصه‌سازی خودکار متن، شبکه‌ی SummaRuNNer است [50]. در این مدل مسئله‌ی خلاصه‌سازی به صورت یک مسئله‌ی دسته‌بندی دنباله‌ای^۱ در نظر گرفته شده است. این مدل از شبکه‌ی عصبی GRU دوطرفه^۲ برای به دست آوردن اطلاعات زمینه‌ای در متن برای تولید مدل استفاده کرده است. همان‌طور که در شکل ۱۸ آمده است، این مدل از سه بخش لایه‌ی کلمات^۳، لایه‌ی جملات^۴ و لایه‌ی دسته‌بندی^۵ تشکیل شده است. در بخش لایه‌ی کلمات، با استفاده از روش word2vec، کلمات به بردارها تبدیل می‌شوند. بردار جمله از گذاشتن بردار کلمات درون یک جمله از یک شبکه‌ی RNN به دست می‌آید. در لایه‌ی جملات نیز یک شبکه‌ی RNN به کار رفته است. در این مدل با استفاده از روابط ۱۳ تا ۱۶ بردار متناسب با هر جمله به دست می‌آید.

$$u_j = \sigma(W_{ux}x_j + W_{uh}h_{j-1} + b_u) \quad (13)$$

$$r_j = \sigma(W_{rx}x_j + W_{rh}h_{j-1} + b_r) \quad (14)$$

¹ Sequence classification

² Bidirectional gated recurrent unit

³ Word layer

⁴ Sentence layer

⁵ Classification layer

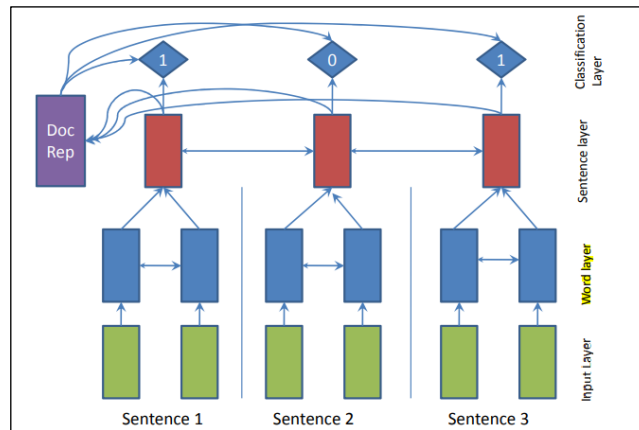
$$h'_j = \tanh(W_{hx}x_j + W_{hh}(r_j \odot h_{j-1}) + b_h) \quad (15)$$

$$h_j = (1 - u_j) \odot h'_j + u_j \odot h_{j-1} \quad (16)$$

در روابط فوق، h_j و x_j بردار مربوط به خروجی بردار مربوط به جمله‌ی z ام است. منظور از U گیت برورسانی، r گیت باز راه‌اندازی و \odot نشان‌دهنده‌ی ضرب مؤلفه به مؤلفه^۱ می‌باشند. بعد از بدست آمدن بردار هر جمله و وارد شدن آن‌ها به یک ترنسفورمر غیرخطی، با استفاده از یک لایه‌ی دسته‌بندی با تابع فعال‌سازی^۲ سیگموئیدی^۳، جملات به دو دسته‌ی ۰ یا ۱ (به معنای حضور یا عدم حضور در خلاصه) تقسیم می‌شوند. همچنین در این مدل، همانطور که در رابطه‌ی ۱۷ آمده است، یک نمایش برای سند، بدست می‌آید. این نمایش سند، در بخش دسته‌بندی به تشخیص برچسب‌ها کمک می‌کند.

$$d = \tanh \left(W_d \cdot \frac{1}{N_d} \sum_{j=1}^{N_d} [h_j^f, h_j^b] + b \right) \quad (17)$$

در رابطه‌ی بالا، h_j^f و h_j^b به ترتیب مقدار بردار عبوری و پیشرو^۴ برای جمله‌ی z ام می‌باشد، که با هم ادغام می‌شوند. همچنین N_d تعداد جملات در سند می‌باشد.



شکل ۱۸: لایه‌های مدل SummaRuNNer [50].

¹ Hadamard product

² Activation function

³ Sigmoid

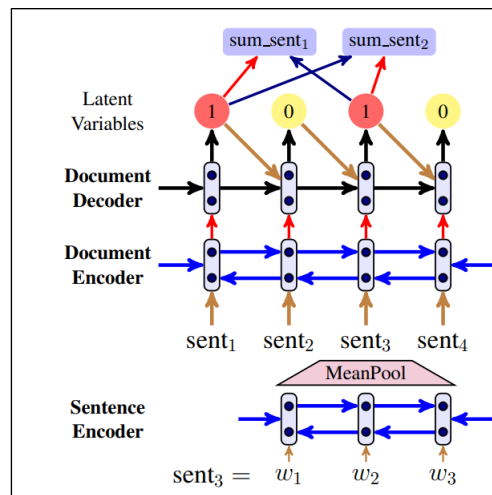
⁴ backward

⁵ forward

از دیگر مدل‌های معروف در این زمینه می‌توان به مدل LatentExtSum نام برد [66]. در مدل LatentExtSum، از مکانیزم نگاشت جملات انتخابی به جملات خلاصه‌ی موجود در مجموعه‌دادگان و با استفاده از تعریف متغیر نهفته^۱، سعی در آن شده است که در فرایند یادگیری، جملات با مفهوم نزدیک به خلاصه‌ی داده شده، استخراج شوند. همان‌طور که در شکل ۱۹ آمده است، این مدل از سه بخش اصلی رمزگذاری جملات^۲، رمزگذاری سند^۳ و رمزگشایی سند^۴ تشکیل شده است. در بخش رمزگذاری جملات، هر جمله با استفاده از یک لایه‌ی LSTM دوطرفه و رابطه‌ی ۱۸ به بردار تبدیل می‌شود. در بخش رمزگذاری سند نیز از یک لایه‌ی LSTM دوطرفه برای یادگیری نمایش زمینه‌ای جملات و در بخش رمزگشایی سند از یک لایه‌ی LSTM برای برچسب‌زنی جملات استفاده شده است. در آخر با استفاده از یک متغیر نهفته که متغیری با مقدار ۰ یا ۱ (به معنای عدم انتخاب یا انتخاب جمله در خلاصه) و سعی در ۱ شدن جملات متن با بیشترین شباهت به خلاصه‌ی انسانی، هستند. به طور ساده این روش با استفاده از روش بیشینه‌سازی شباهت^۵، سعی دارد جملات شبیه به خلاصه‌ی انسانی استخراج کند.

$$v_i = \frac{1}{|S_i|} \sum_j h_j^i \quad (18)$$

در رابطه‌ی فوق منظور از $|S_i|$ اندازه‌ی جمله‌ی i ام می‌باشد و منظور از h_j^i بردار حالت پنهان^۶ j ام LSTM برای جمله‌ی i ام می‌باشد.



شکل ۱۹: معماری کلی LatentExtSum [66].

¹ Latent variables

² Sentence encoder

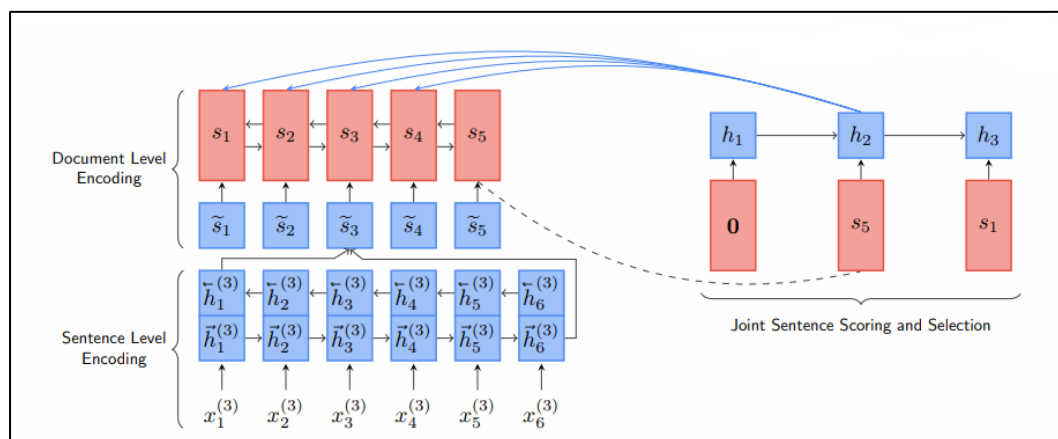
³ Document encoder

⁴ Document decoder

⁵ Maximum likelihood

⁶ Hidden state

در مدل NeuSum [56] از دو نوع رمزگذاری در سطح جمله و در سطح سند استفاده می‌کند و چهارچوبی برای چگونگی امتیازدهی و انتخاب جملات ارائه می‌دهد. همانطور که در شکل ۲۰ آمده است، برای جملات یک بار با استفاده از رمزگذاری سطح جمله و بار دیگر با استفاده از رمزگذاری سطح سند، به صورت سلسله مراتبی بردار ویژگی بدست می‌آید. در بخش رمزگذاری سطح جمله، کلمات استخراج شده و با استفاده از روش‌های جانشانی کلمات (برای مثال glove)، بردار ویژگی هر کلمه بدست می‌آید. با استفاده از شبکه‌ی عصبی بازگشتی GRU دو طرف (Bi-GRU)، ویژگی‌های زمینه‌ای در سطح هر جمله استخراج می‌شود. همچنین در سطح سند، دنباله‌ای از این ویژگی‌های استخراج شده در سطح جملات، داخل Bi-GRU دیگر شده و به ازای هر جمله ویژگی‌های زمینه‌ای سطح سند استخراج می‌گردد. در آخر نیز ویژگی‌های سطح جملات و سطح سند با استفاده از یک MLP دو لایه امتیازدهی شده و انتخاب می‌شوند.



شکل ۲۰: استخراج ویژگی‌های جملات به صورت سلسله مراتبی در دو سطح جمله و سند [56].

مدل شبکه‌ی عصبی از پیش آموزش داده شده‌ی BERT، به منظور جانشانی کلمات با استفاده از یادگیری مدل زبانی برای نمایش ویژگی‌های زمینه‌ای بر روی یک پیکره‌ی متنی مقیاس بزرگ، به وجود آمده است. این شبکه، نمایش کلمات و جملات را با استفاده از یک ترنسفورمر بسیار بزرگ با هم ادغام می‌کند و توانسته است در مسائل پردازش متن دست‌آوردهای بسیار زیادی داشته باشد.

مدل BertSumExt [47] یک مدل خلاصه‌سازی متن با رویکرد استخراجی می‌باشد. این مدل با استفاده از مدل از پیش آموزش دیده‌ی BERT و قراردادن چند لایه‌ی ترنسفورمر جمله‌ای در بالای رمزگذار به منظور در نظر گرفتن ویژگی‌های سطح سند، پیشرفت قابل توجهی در حوزه‌ی خلاصه‌سازی خودکار متن به دست آورده است. شکل ۲۱، معماری مدل BERT اصلی و BERT برای خلاصه‌سازی متن را به نمایش گذاشته شده است. همان‌طور که در شکل ۲۱ آمده است، در معماری BERT اصلی، جملات با توکن‌های [SEP] جدا شده و در ابتدای هر متن یا سند توکن [CLS] قرار می‌گیرد. در حالی که در معماری BERT برای خلاصه‌سازی، به منظور نمایش مستقلانه‌ی جملات و یادگیری نمایش

جملات مخصوص خلاصه‌سازی، توکن [cls] در ابتدای هر جمله می‌آید. در هر دو معماری برای هر کلمه سه جانشانی در نظر گرفته شده است. اولین آن‌ها جانشانی توکن^۱، به منظور نمایش معنای هر کلمه، دومین جانشانی که جانشانی بخش^۲ است، به منظور تمایز دادن بین جملات و سومین آن‌ها، جانشانی موقعیت^۳ به منظور نمایش موقعیت هر توکن است. این سه جانشانی با هم ادغام شده و به لایه‌های ترنسفورمر داده می‌شود و در آخر برای هر کلمه در جمله بردار مناسبی به دست می‌آید. در مدل پیشنهادی، بردار t_i که نمایش دهنده‌ی جمله‌ی i ام است، خروجی لایه‌ی آخر شبکه می‌باشد. خروجی هر لایه از شبکه به صورت روابط ۱۹ و ۲۰ است.

$$\tilde{h}^l = LN(h^{l-1} + MHAtt(h^{l-1})) \quad (19)$$

$$h^l = LN(\tilde{h}^l + FFN(\tilde{h}^l)) \quad (20)$$

در روابط فوق، LN نمایش دهنده‌ی لایه‌ی نرمال‌سازی^۴، MHAtt به منظور عملگر توجه چند سره و h^l نمایش دهنده‌ی خروجی لایه‌ی l است. همچنین مقدار h^0 برابر با PosEmb(T) می‌باشد که در آن، T به منظور خروجی BERT برای خلاصه‌سازی و PosEmb نیز یک تابع جانشانی کننده‌ی مکانی سینوسی^۵ است. همچنین در بخش دسته‌بند، مطابق با رابطه‌ی ۲۱ از یک لایه‌ی سیگموئیدی استفاده شده است.

$$\hat{y}_i = \sigma(W_0 h_i^L + b_0) \quad (21)$$

در روابط فوق منظور از h_i^L بردار خروجی جمله‌ی i ام با شبکه‌ی L لایه (که در آزمایشات تعداد ۲ لایه مناسب دیده شده است)، می‌باشد. W_0 و b_0 مؤلفه‌های قابل آموزش هستند.

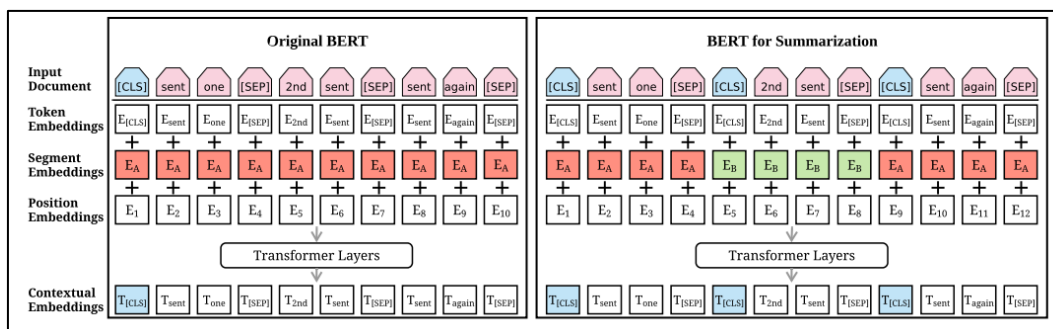
¹ Token embedding

² Segment embedding

³ Position embedding

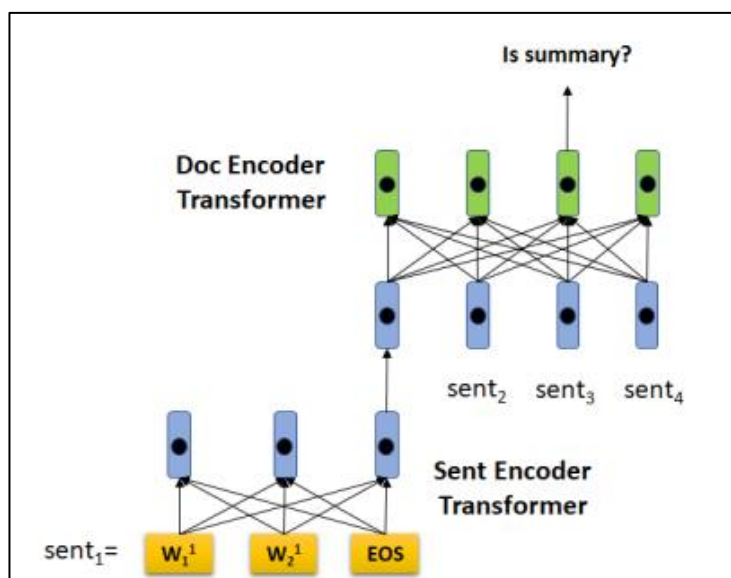
⁴ Normalization layer

⁵ Sinusoid positional embeddings



شکل ۲۱: مدل BERT اصلی و BERT برای خلاصه‌سازی متن [47].

پس از پیشرفت‌های مهم مدل BERT در زمینه خلاصه‌سازی متن، مقالات زیادی تلاش کرده‌اند تا بر روی این دستاوردها ادامه دهند. یکی از مدل‌های پیشنهادی که در همین راستا استفاده از مدل HIBERT می‌باشد [67]. مدل‌های خودکار خلاصه‌سازی عصبی معمولاً از یک کدگذار سلسله مراتبی برای کدگذاری اسناد استفاده می‌کنند و با استفاده از برجسب‌های سطح جمله آموزش داده می‌شوند. HIBERT یک مدل برای پردازش و خلاصه‌سازی اطلاعات در سطح سند است. این مدل از معماری ترنسفرمر برای نمایش و کدگذاری سند استفاده می‌کند و با ترکیب دو کدگذار، یکی برای جملات و دیگری برای سند، به صورت سلسله مراتبی ساخته شده است. هر سند شامل جملات است که با استفاده از کدگذار جمله به بردارهای مربوط به جملات تبدیل می‌شوند. سپس با استفاده از یک کدگذار سند بردار نهایی سند به دست می‌آید. بیشتر مدل‌های رمزگذار، کلمه را متناسب با کلمات قبل و بعد آن در جمله پیش‌بینی می‌کنند، در حالی که مدل از پیش‌آموزش دیده‌ی HIBERT سعی دارد تا یک جمله را با توجه به جملات قبل و بعد آن پیش‌بینی کند. همانطور که در شکل ۲۲ آمده است، در این مدل از دو رمزگذار استفاده شده است. رمزگذار جمله، ویژگی‌های مربوط به هر جمله و رمزگذار سند، ویژگی‌های مربوط به زمینه‌ی متن را استخراج می‌کند. با استفاده از این مدل رمزگذارها، سند و جملات به صورت سلسله مراتبی، رمزگذاری می‌شوند، به طوری که از یک ترنسفرمر کلمات هر جمله عبور داده شده و ویژگی آخر آن‌ها به عنوان ویژگی جمله در نظر گرفته می‌شود. با استفاده از یک ترنسفرمر دیگر، ویژگی جملات در سطح سند استخراج می‌شوند. در آخر با استفاده از یک لایه‌ی دسته‌بند مشخص می‌شود چه جملاتی در خلاصه حضور پیدا خواهند کرد یا خیر [67].



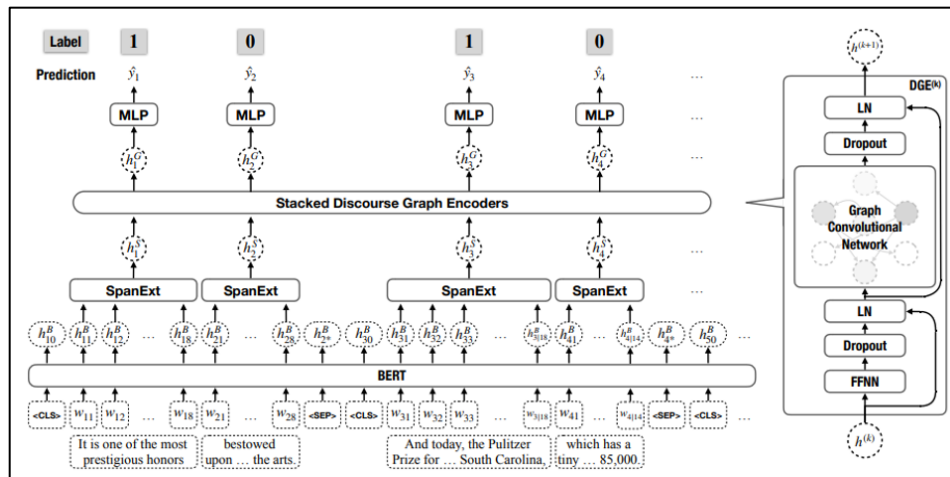
شکل ۲۲: کدگذاری سلسله مراتبی جملات در روش HIBERT [67].

۳-۴-۲- مدل‌های مبتنی بر شبکه‌های عصبی گراف

خلاصه‌سازی خودکار متن با استفاده از شبکه‌های عصبی گراف یک حوزه پژوهشی مهم است که در آن، سعی بر این دارد که با بهره‌گیری از قابلیت‌های گرافی مدل‌های عصبی، به‌طور هوشمندانه‌تر و جامع‌تر محتوای یک متن را خلاصه کند. در این رویکرد، معمولاً گره‌های گراف به اجزای مهم متن مثل جملات یا کلمات شکسته می‌شوند و یال‌ها نشان‌دهنده ارتباطات میان اجزا هستند. مدل‌های گراف عصبی می‌توانند با توجه به این ارتباطات و ساختار گراف، خلاصه متنی تولید کنند که اطلاعات کلیدی را حفظ کرده و مفهوم متن اصلی را به‌دست آورند. این رویکرد جدید نسبت به روش‌های سنتی خلاصه‌سازی متن، که معمولاً از روش‌های آماری یا مبتنی بر قوانین استفاده می‌کنند، امکانات بیشتری برای درک متن و تولید خلاصه با کیفیت‌تر فراهم می‌کند. این مدل‌ها معمولاً از معماری‌هایی مانند شبکه‌های توجه بهره‌می‌برند تا به‌طور هوشمندانه به اجزا و ارتباطات مهم توجه کنند. با پیشرفت تکنولوژی گرافی و توانمندی‌های شبکه‌های عصبی، انتظار می‌رود که این رویکرد در آینده بیشتر به کار گرفته شود و به بهبود مدل‌های خودکار خلاصه‌سازی متن کمک کند [8]. [54].

همان‌طور که پیش‌تر ذکر شد، یکی از بزرگ‌ترین چالش‌های موجود در سیستم‌های خلاصه‌سازی، چگونگی در نظر گرفتن ویژگی‌های زمینه‌ای با فاصله از هم می‌باشد. با وجود این که برخی مدل‌ها مثل مدل BERT تا حدی توانست این موضوع را در نظر بگیرد، ولی برخی ضعف‌ها در این زمینه وجود دارد؛ حتی وجود LSTM‌ها و ترنسفورمرها در بالای BERT نیز بهبودی حاصل

نمی‌کند. مدل DiscoBERT [68] با استفاده از گراف‌هایی مثل RST^۱ و مرجع^۲ و با یادگیری شبکه‌های عصبی پیچشی گراف (GCN) توانسته است ویژگی‌های زمینه‌ای طولانی بین جملات را در نظر بگیرد و باعث بهبود فشرده‌سازی و کاهش تکرار شود. در مدل DiscoBERT، ابتدا از متن ورودی یکی از دو گراف RST یا مرجع ساخته می‌شود. گراف RST با استفاده از درخت تجزیه‌ی RST ساخته شده و گراف مرجع با استفاده از همبستگی‌های بین موجودیت‌های متن ورودی^۳ تولید می‌شود. همان‌طور که در شکل ۲۳ آمده است، معماری مدل DiscoBERT از BERT به‌عنوان رمزگذاری سند استفاده می‌کند، سپس با استفاده از واحدهایی به نام SpanExt واحدهایی به نام EDU^۴ که کوچک‌ترین واحد خلاصه‌سازی هستند، استخراج می‌شود. بعد از استخراج EDU ها، گراف موردنظر ساخته شده و با استفاده از شبکه‌ی GCN، بردارهای مخصوص هر EDU استخراج و وارد شبکه‌ی عصبی MLP شده و دسته‌بندی می‌شوند.



شکل ۲۳: معماری کلی از شبکه‌ی DiscoBERT [68].

مدل GraphSumTextRank [69] یکی دیگر از مدل‌های مبتنی بر گراف است. در این مقاله یک روش خلاصه‌سازی متن مبتنی بر گراف توضیح داده شده است. در زمان‌های اخیر، به دلیل رشد سریع اینترنت، دسترسی به میزان عظیمی از اطلاعات به یک چالش بزرگ تبدیل شده است. برای مدیریت این حجم زیاد از اطلاعات، نیاز به روش‌ها و ابزارهای کارآمد و موثر مثل روش PageRank وجود دارد. PageRank یک الگوریتم مهم در حوزه موتورهای جستجو و ارتباطات وب است که توسط

¹ Rhetorical Structure Theory

² Coreference

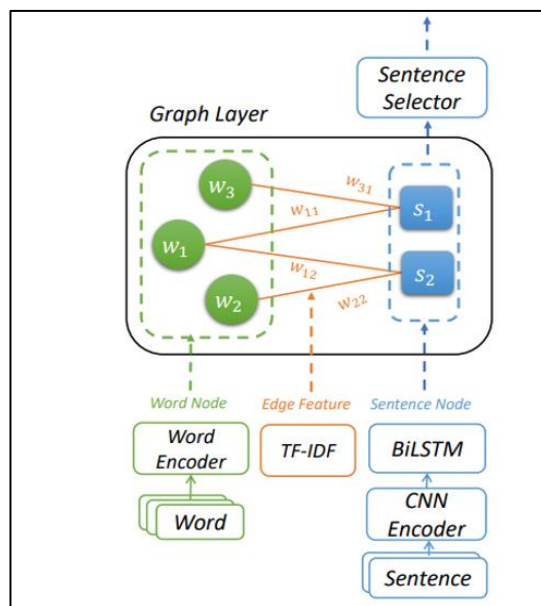
³ coreference mentions based

⁴ Elementary Discourse Unit

لری پیج و سرگی برین در دانشگاه استنفورد توسعه یافته است. این الگوریتم برای ارزیابی اهمیت یک صفحه وب استفاده می‌شود و ایده اصلی آن بر اساس ساختار گرافی اینترنتی است. در PageRank، هر صفحه به عنوان یک گره در گراف در نظر گرفته می‌شود و وزن هر گره بر اساس تعداد وزن‌دهی شده یال‌های وارد به آن صفحه محاسبه می‌شود. به عبارت دیگر، یک صفحه ارزش بیشتری دارد اگر صفحات با ارتباط به آن ارتباطات بیشتری داشته باشند. روش معرفی شده در این مقاله از TextRank اصلاح شده استفاده می‌کند که بر اساس مفهوم PageRank تعریف شده برای هر صفحه در وب‌سایت‌ها است. این روش یک گراف با جملات به عنوان گره‌ها و شباهت بین دو جمله به عنوان وزن یال بین آن‌ها ایجاد می‌کند. سپس گراف ساخته شده و به خوشه‌های مختلف تقسیم می‌شود با فرض اینکه جملات داخل یک خوشه به یکدیگر شبیه هستند و جملات خوشه‌های مختلف نشان‌دهنده تفاوت آن‌ها هستند. ارزیابی عملکرد این تکنیک خلاصه‌سازی پیشنهادی نشان می‌دهد که این روش کارآمدی دارد.

یکی دیگر از مدل‌هایی که سعی کرده است تا وابستگی‌های با فاصله‌ی زیاد در متن را در نظر بگیرد، مدل HSG [51] می‌باشد. در این مدل با طراحی گراف ناهمگن، توانسته است پیشرفت قابل توجهی در زمینه‌ی خلاصه‌سازی متن به دست آورد. به منظور ترکیب ویژگی‌های جملات مختلف و همچنین کلمات مختلف، جملات و کلمات به‌عنوان گره‌هایی از یک گراف پیشنهاد شده است. همان‌طور که در شکل ۲۴ دیده می‌شود، در این روش کلمات و جملات به‌عنوان گره‌های گراف و بین هر گره‌ی کلمه با گره‌ی جمله به‌شرط وجود آن کلمه در جمله یالی قرار می‌گیرد. همچنین وزن هر یال بر اساس معیار TF-IDF و میزان تکرار کلمه در جمله محاسبه می‌شود. بردار موجود در گره‌های کلمه به وسیله‌ی یک شبیه‌ساز کلمه و بردار موجود در گره‌های جملات به وسیله‌ی یک رمزگذار پیچشی^۱ و شبکه‌ی BiLSTM، به دست می‌آیند.

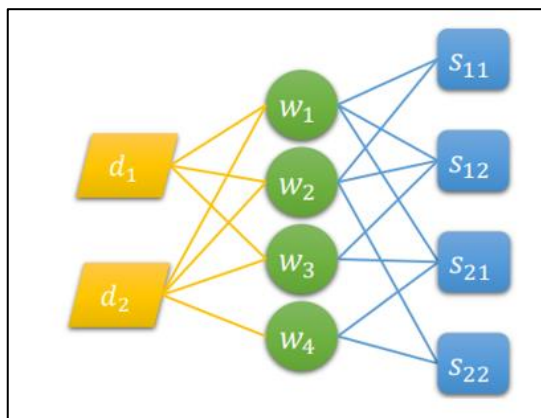
¹ CNN encoder (convolutional neural network encoder)



شکل ۲۴: معماری کلی وساخت گراف در مدل HSG [51].

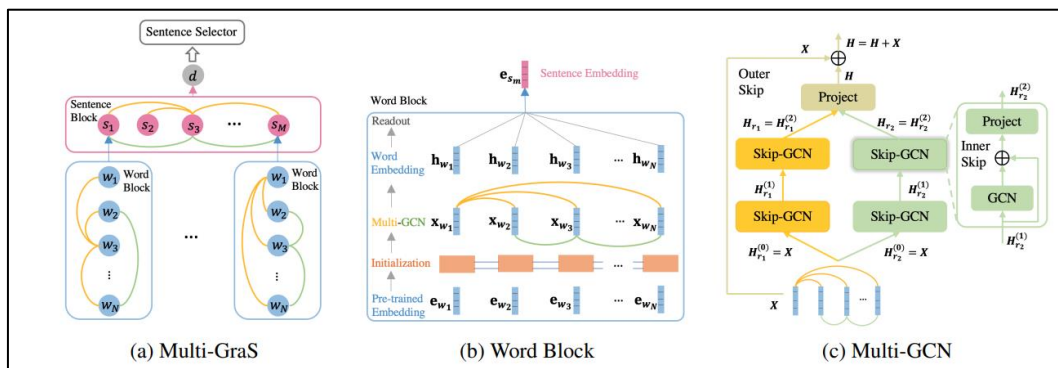
در این روش، برای بخش خلاصه‌سازی چند سندی نیز راه حلی وجود دارد که باعث به‌دست‌آمدن نتایج مفید شده است. همان‌طور که در شکل ۲۵ آمده است، در روش HDSG به مشابه گراف ذکر شده در قسمت قبل، گره‌های جملات و کلمات به دست می‌آید و گره‌هایی برای هر سند در نظر گرفته می‌شود. بردار درون هر گرهی سند و نحوه‌ی اتصال یال‌ها مشابه گره‌ی جملات می‌باشد.

در قسمت آموزش گراف از ساختار توجه چند سره و گراف توجه و ساختار اتصال residual استفاده می‌شود. بعد از اعمال الگوریتم پیام‌رسانی و جابه‌جایی اطلاعات بین گره‌ها، در بخش انتخاب جملات، به گره‌های جملات امتیازدهی شده و گره‌های با امتیاز بالا انتخاب می‌شوند.



شکل ۲۵: شمای کلی گراف در مدل HDSG [51].

یکی دیگر از روش‌های گرافی که توانسته است نتایج خوبی را به دست آورد، مدل Multi-GraS [70] می‌باشد. همان‌طور که در شکل ?? آمده است، این مدل از سه بخش اصلی بلوک کلمات، بلوک جملات و بخش انتخاب جملات تشکیل شده است. در بخش بلوک کلمات، گرافی شامل گره‌های کلمات یک جمله، تشکیل می‌شود. در این گراف یال‌ها بر اساس بر اساس چندین روابط از جمله روابط نحوی استخراج شده با تجزیه‌گر وابستگی^۱، بین دو گرهی کلمه در هر جمله به وجود می‌آیند. در بخش بلوک کلمات، هر کلمه به وسیله‌ی جانشانی کلمات، تبدیل به بردار می‌شود. با استفاده از مکانیزم Multi-GCN که در شکل ۲۶ آمده است، عملیات آموزش و تبادل اطلاعات بر روی گراف صورت می‌گیرد. بردار خروجی هر کلمه با هم ادغام شده و بردار جمله‌ی مربوطه تشکیل می‌شود. در بلوک جملات، هر جمله به‌عنوان یک گره در نظر گرفته شده و یال‌های این گراف بر اساس میزان شباهت بین دو جمله و تعداد کلمات مشترک بین دو جمله به وجود می‌آیند. در این بخش نیز با استفاده از مکانیزم Multi-GCN عملیات آموزش و تبادل اطلاعات بر روی گراف صورت می‌گیرد. در ساختار GCN-Multi از چند شبکه‌ی skip-GCN و ادغام نتایج آن‌ها استفاده شده است. در بخش انتخاب جملات، با استفاده از یک شبکه‌ی عصبی پیشرو جملات امتیازدهی شده و جملات با امتیاز بالاتر استخراج می‌شوند.



شکل ۲۶: معماری شبکه Multi-GraS [70].

با وجود عبارات تکراری بین جملات استخراج شده، مدل کردن دقیق جملات با روش‌های عمومی نظارت شده مشکل است. مدل HAHSUM^۲ [71] به خوبی اطلاعات مختلف را مدل می‌کند و با استفاده از شبکه‌های عصبی گراف، وابستگی‌های تکراری بین جملات را برجسته می‌کند. در این روش یک گراف ناهمگن که حاوی سطوح اطلاعات چندگانه برای نمایش یک سند است ساخته

¹ Dependency parser

² Hierarchical Attentive Heterogeneous Graph for Text Summarization

می‌شود. در این گراف از سه نوع گره کلمات، جملات و موجودیت‌های نامدار^۱ استفاده شده است. موجودیت‌های نامدار و کلمات متوالی داخل یک جمله به هم وصل می‌شوند. کلمات و موجودیت‌های یک جمله به گره‌ی آن جمله به صورت یال یک طرفه وصل می‌شوند. همچنین گره‌های موجودیت‌های نامدار یکسان نیز به هم متصل می‌شوند. در آخر جملاتی که دارای trigram یکسان باشند نیز به هم متصل می‌شوند. با استفاده از مدل از پیش آموزش دیده‌ی ALBERT ویژگی‌های گره‌های کلمات و جملات بدست می‌آید. با استفاده از شبکه‌ی عصبی GAT و به وسیله‌ی کاهش اطلاعات تکراری در فرایند انتقال پیام باعث می‌شود تا جملات با محتوای تکراری امتیاز کمتری بدست آورند و خلاصه از پراکندگی اطلاعات بهتری برخوردار باشد.

استفاده از گراف‌ها موفقیت‌های زیادی در مسئله‌ی خلاصه‌سازی متن داشته است. همچنین بسیاری از روش‌های مبتنی بر عنوان نیز از روش‌های گرافی استفاده می‌کنند. در ادامه به بررسی روش‌های مبتنی بر عنوان نیز بررسی شده است.

۳-۴-۳- روش‌های مبتنی بر عنوان

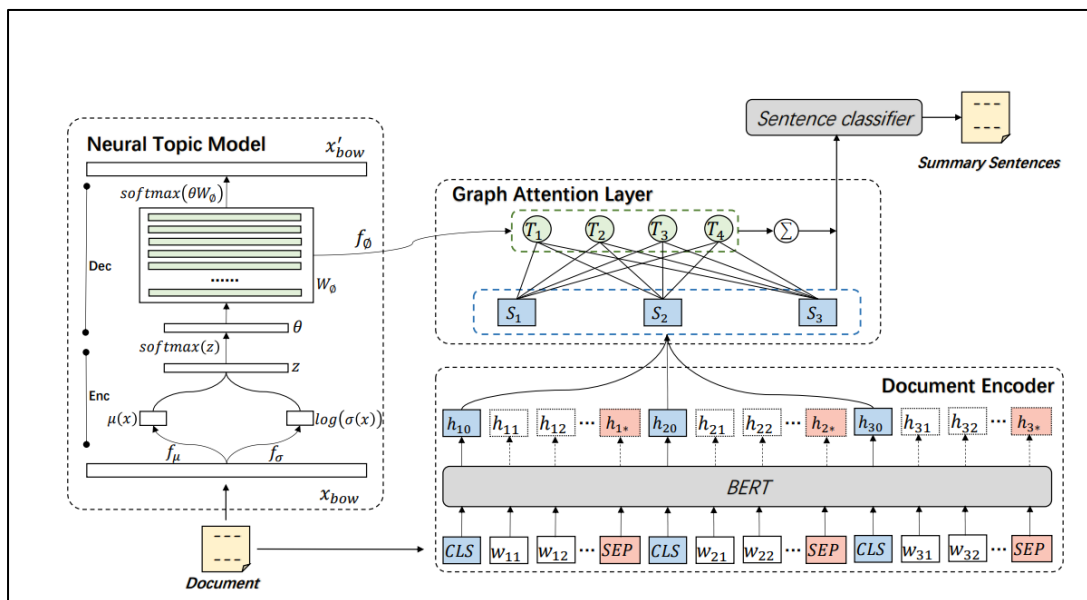
خلاصه‌سازی متن مبتنی بر عنوان یک روش مؤثر در پردازش زبان طبیعی است که بر اساس تشخیص موضوعات اصلی در متن، سعی در تولید خلاصه مطالب می‌کند. در این روش، ابتدا موضوعات کلیدی و عناوین مهم متن شناسایی می‌شوند، سپس با ترکیب و خلاصه‌سازی جملات مرتبط با این موضوعات، یک خلاصه مفید و مختصر از مطلب اصلی تولید می‌شود. استفاده از خلاصه‌سازی مبتنی بر عنوان به تحلیل محتوای متن و شناسایی مفاهیم اصلی می‌پردازد، که این امر می‌تواند در فهم سریع‌تر و بهتر متون به کار گیرد. این روش مخصوصاً در وظایف مانند خلاصه‌سازی اخبار، مقالات و مطالب طولانی کمک بزرگی به کاربران جهت به دست آوردن یک نگاه کلی از محتوا دارد.

خلاصه‌سازی متون با مقیاس بزرگ یکی از اصلی‌ترین چالش‌های مسئله‌ی خلاصه‌سازی متن است. مدل TopicGraphSum [72] با در نظر گرفتن روابط بین جملات در متن و عنصر مهم عنوان متن، توانسته است در متون مقیاس بزرگ، موفق ظاهر شود. مدل ارائه شده، حتی نتایج خوبی در متون کوتاه نیز داشته است. در این مدل با استفاده از مدل عصبی مبتنی بر عنوان (NTM^۲)، عنوان هر متن استخراج می‌شود. همان‌طور که در شکل ۲۷ آمده است، این مدل شامل بخش رمزگذاری سند می‌باشد. در این بخش سند به وسیله‌ی BERT رمزگذاری شده و کلمات به بردار تبدیل

¹ Named entity

² Neural topic model

می‌شود. علائم CLS و SEP برای شناسایی جملات در این بخش به کار می‌روند. در بخش NTM، استخراج بردار موضوع از سند ورودی با استفاده از پروسه‌ی رمزگذاری - رمزگشایی^۱ صورت می‌گیرد. ورودی این بخش، نمایش کیسه‌ی کلمات از سند است و این بخش آموزش می‌بیند تا بردار ویژگی مناسبی برای عنوان، استخراج شود. در قسمت بعد، یک گراف ناهمگن، شامل گره‌های جملات و عنوان‌ها ساخته شده و همه‌ی جملات یک سند به گره‌ی عنوان وصل می‌شوند. به همین دلیل، همه‌ی جملات با یک مسیر به طول ۲ به هم متصل هستند (برای ارتباط بین جملات در یک موضوع و سند). با استفاده از مکانیزم توجه گراف، اطلاعات اشتراک گذشته شده و روابط معنایی بین جملات یاد گرفته و در آخر توسط یک دسته‌بند جملات، گره‌های مناسب برای خلاصه استخراج می‌شود.



شکل ۲۷: معماری مدل TopicGraphSum [72].

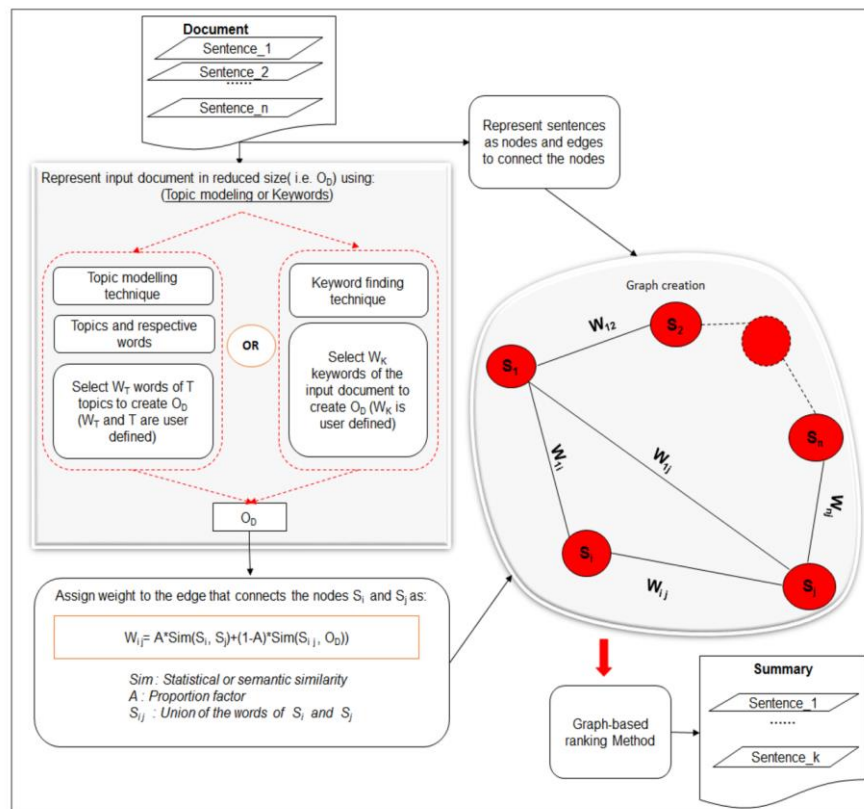
در تکنیک‌های خلاصه‌سازی متن مبتنی بر گراف، وزن نسبت داده شده به یال‌های گراف پارامتر حیاتی برای رتبه‌بندی جملات است. این وزن‌ها بر اساس شباهت بین جملات (گره‌ها) تعیین می‌شوند. بیشتر تکنیک‌های مبتنی بر گراف از اندازه‌گیری شباهت کلمات مشترک برای نسبت دادن وزن استفاده می‌کنند. در مدل ارائه شده توسط بلوال و همکاران [73]، یک تکنیک جدید خلاصه‌سازی مبتنی بر گراف پیشنهاد شده است که علاوه بر در نظر گرفتن شباهت بین جملات متن، شباهت بین جملات و متن کلی ورودی را نیز مد نظر قرار می‌دهد. در حین نسبت دادن وزن به یال‌های گراف، دو ویژگی در نظر گرفته می‌شود. ویژگی اول شباهت بین گره‌هایی است که یال‌های

¹ Encoding-decoding process

گراف را تشکیل می‌دهند. ویژگی دوم وزن داده شده به یک مؤلفه است که نشان‌دهنده این است که چقدر یک یال خاص با موضوعات کلی متن ورودی مشابه است، که در این مدل، موضوع را در آن ادغام می‌کند. به همراه این تغییرات، از اندازه‌گیری معنایی برای یافتن شباهت بین گره‌ها استفاده می‌شود. روند کلی این روش در شکل ۲۸ به نمایش گذاشته شده است. در این روش ابتدا گراف کاملی با گره‌های جملات ساخته می‌شود، همچنین با استفاده از روش‌های متداول استخراج عنوان یا استخراج کلمات کلیدی (روش‌هایی مثل LDA topic modeling)، مجموعه از عبارات اصلی متن استخراج می‌شود. با استفاده از معادله ۲۲ وزن بین هر گره مشخص می‌شود. به منظور مرتب سازی جملات بر اساس معیار اهمیت از الگوریتم graph ranking استفاده می‌شود. جملات با رتبه‌ی بالاتر به منظور تولید خلاصه انتخاب می‌شوند.

$$W(S_i, S_j) = A * Sim(S_i, S_j) + (1 - A) * Sim(S_{ij}, O_D) \quad (22)$$

در عبارت ریاضی بالا، A ضریب اهمیت شباهت بین دو جمله می‌باشد. در قسمت دوم شباهت کلمات مشترک دو جمله با مجموعه کلیدواژه‌ها محاسبه می‌شود. توابع شباهت بین دو مجموعه کلمه یا دو جمله با استفاده از روش‌هایی مثل شباهت کسینوسی، شباهت اقلیدسی و شباهت جاکارد محاسبه می‌شود.



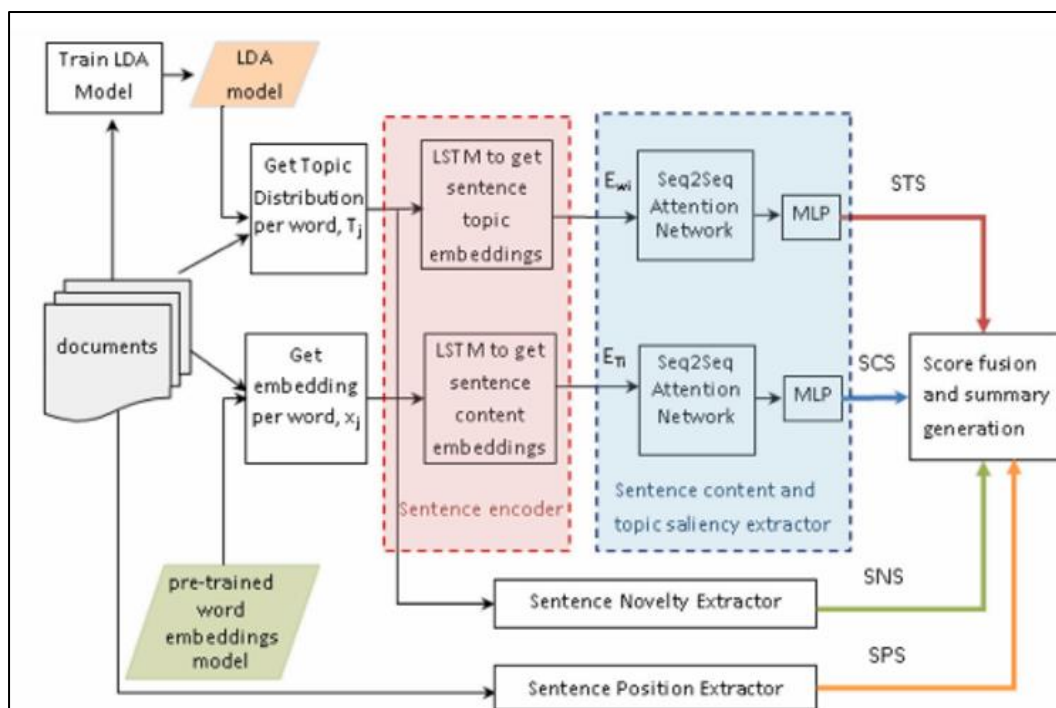
شکل ۲۸: نمای کلی از روش ارائه شده توسط بلوال و همکاران و استفاده از گراف و مجموعه کلیدواژه‌ها یا عناوین برای خلاصه‌سازی متن [73].

در مدل DeepSum [74]، روش نوآورانه‌ای برای خلاصه‌سازی متن استخراجی از اسناد تکی ارائه شده است. این روش بر اساس مدل‌های موضوع و جاسازی کلمات برای خلاصه‌سازی متن‌ها استفاده می‌کند. روش‌های اخیر بر اساس شبکه‌های دنباله‌ای ناتوان در درک مفاهیم با طول متن بلند اسناد هستند که در بردارهای موضوع اسناد جاسازی شده‌اند. در این مدل، هدف استفاده از اطلاعات پنهان در اسناد از طریق بردارهای موضوع و شبکه‌های دنباله‌ای است تا کیفیت و دقت متن خلاصه شده را بهبود بخشد.

هر جمله از طریق دو شبکه عصبی بازگشتی مختلف بر اساس توزیع احتمالی موضوع و جاسازی کلمات ساخته شده، سپس یک شبکه RNN به هر کدام از این جملات اعمال می‌شود. خروجی‌های کدگذار و کدگشا در شبکه دنباله‌ای پس از وزن‌دهی با استفاده از یک مکانیزم توجه ترکیب شده و به وسیله یک شبکه ورودی چندلایه به امتیاز تبدیل می‌شوند. امتیاز حاصل از مدل موضوع^۱ (STS) به عنوان امتیاز موضوع جمله و امتیاز حاصل از جاسازی کلمات به عنوان امتیاز محتوای جمله (SCS)

¹ Sentence topic score

¹ اشاره دارد. علاوه بر این، امتیاز جدید نوآوری جمله یا عدم تکراری بودن^۲ (SNS) و امتیاز موقعیت جمله^۳ (SPS) ارائه شده و انجام ادغام وزن دار چهار امتیاز برای هر جمله در سند جهت محاسبه امتیاز نهایی جمله (FSS) انجام می‌شود. برای محاسبه‌ی امتیاز SPS، میزان جایگاه نسبی جمله به کل جملات می‌باشد به طوری که جملات ابتدایی امتیاز بالاتری می‌گیرند. همچنین امتیاز SPS نیز میانگین میزان شباهت کسینوسی جمله مورد نظر با جملات قبل از آن می‌باشد. شکل ۲۹ معماری کلی این مدل را نمایش می‌دهد.



شکل ۲۹: معماری کلی روش DeepSum [74].

¹ Sentence context score

² Sentence novelty score

³ Sentence position score

۳-۵- جمع‌بندی

در این بخش به رویکردها و روش‌های رایج در مسئله‌ی خلاصه‌سازی استخراجی متن پرداخته شد. ابتدا مسئله‌ی خلاصه‌سازی استخراجی متن به صورت یک مسئله‌ی دسته‌بندی دودویی تعریف شد بدین صورت که آیا هر جمله از متن ورودی در جملات انتخاب شده در خلاصه هست یا خیر. استخراج برچسب‌ها از خلاصه‌ی انتزاعی به روش oracle انجام می‌شود. در ادامه، روش‌های موجود در خلاصه‌سازی استخراجی متن به سه دسته‌ی کلی روش‌های سنتی، یادگیری ماشین و یادگیری عمیق تقسیم بندی شده و برای هر بخش مدل‌ها و روش‌های معروف ذکر شدند. در روش‌های سنتی معمولاً اهمیت مفاهیم TF-IDF، موقعیت مکانی و شباهت جملات با سند مورد بررسی قرار می‌گیرد. با رشد الگوریتم‌های یادگیری ماشین مخصوصاً در حوزه‌های دسته‌بندی و یادگیری تقویتی، پیشرفت‌های قابل ملاحظه‌ای در مسئله‌ی خلاصه‌سازی خودکار متن اتفاق افتاد. SVM، KNN و Naïve Bayes از مدل‌های مؤثر برای دسته‌بندی در مسئله‌ی خلاصه‌سازی هستند. علارغم دست‌آوردهای زیادی یادگیری عمیق در حوزه‌ی خلاصه‌سازی، این روش‌ها همچنان مشکلاتی از قبیل توانایی ضعیف در درک معانی عمیق متون و تولید خلاصه با پوشش بهتر متن، داشتند. یادگیری عمیق در حوزه‌ی خلاصه‌سازی متن توانست برخی نواقص روش‌های سابق را رفع کند و به نتایج بسیار خوبی برسد. در این بخش مدل‌های یادگیری عمیق در خلاصه‌سازی متن با توجه به ایده‌ی اصلی هر مدل، به سه بخش مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی، شبکه‌های عصبی گراف و روش‌های مبتنی بر عنوان، تقسیم شدند. در روش‌های مبتنی بر شبکه‌های عصبی بازگشتی، معمولاً جملات دنباله‌ای از کلمات و سندها دنباله‌ای از جملات در نظر گرفته شده و با استفاده از شبکه‌های عصبی بازگشتی نمایش معنایی از جملات در سند به دست آمده و با استفاده از یک روش انتخابی، جملات خلاصه استخراج می‌گردد. در مسئله‌ی خلاصه‌سازی متن، انواع روابط می‌تواند کمک کند. یکی از روش‌هایی که می‌تواند این روابط را مدل کند و مسائلی از قبیل در نظر گرفتن ویژگی‌های زمینه‌ای با فاصله از هم را حل کند، نمایش گراف برای این استخراج خلاصه از متن می‌باشد. استفاده از گراف و شبکه‌های عصبی گراف باعث می‌شود اطلاعات مختلف با توجه به روابط گره‌ها به اشتراک گذاشته شده و تصمیم‌گیری بهتری اتخاذ شود. همچنین در رویکرد استخراجی متن، دانستن کلیدواژه‌ها یا عناوین در متن، می‌تواند جملات با اهمیت بهتر را استخراج نمود و اطلاعات گوناگون با محوریت عناوین مختلف استخراج کرد.

در ادامه، با توجه به روش‌های گفته شده در ادبیات موضوع و پیشینه پژوهش، سعی شده است مدل‌های مؤثری ارائه شوند و نتایج آن‌ها با مدل‌های معرفی شده در پژوهش‌های قبلی، مورد مقایسه قرار گیرد.

فصل چهارم

روش پیشنهادی

۴-۱- مقدمه

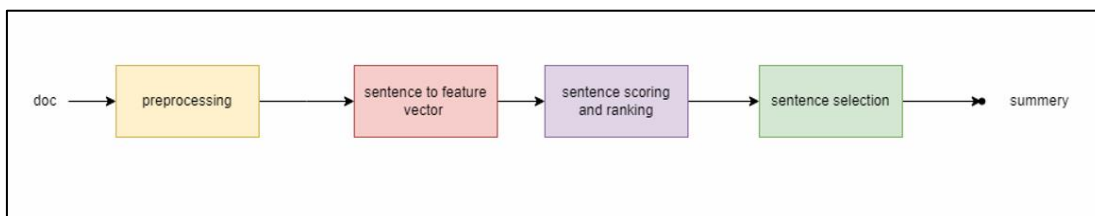
در فصل‌های گذشته، مبانی لازم برای پیاده‌سازی یک سیستم خودکار خلاصه‌سازی متن مورد بحث قرار گرفت. همانطور که پیش‌تر ذکر شد، می‌توان مسئله خلاصه‌سازی متن را با استفاده از رویکرد استخراجی به عنوان یک مسئله دسته‌بندی دوتایی تعریف کرد. در این روش‌ها، بردار ویژگی مناسبی از جملات سند استخراج شده و با استفاده از یک دسته‌بند، جملات به دو دسته خلاصه و غیر خلاصه تقسیم می‌شوند.

خصوصیات متنوعی از کلمات و جملات در خلاصه‌سازی متن نقش دارند. همچنین، روابط موجود بین کلمات و جملات می‌توانند ویژگی‌های مهمی را در این حوزه برجسته کنند. در این پژوهش، تلاش شده است تا با بهره‌گیری از روابط متنوع موجود در یک سند و از دیدگاه‌های مختلف، مدل‌هایی برای استخراج خلاصه ارائه شود. اطلاعات زمینه‌ای، اطلاعات معنایی و نحوی در سطوح مختلف می‌توانند به ساخت یک خلاصه مناسب و مفید کمک کنند. به همین منظور، در این پژوهش،

معماری‌های متنوعی با استفاده از ساختارهای داده‌ای گوناگون، به ویژه گراف، معرفی شده است. این فصل به بررسی جزئیات مختلف این معماری‌ها، همراه با مزایا و نقاط ضعف آن‌ها می‌پردازد. در پایان، جزئیات مدل نهایی پیشنهادی به همراه چگونگی آموزش آن به دقت شرح داده می‌شوند.

۴-۱ - ساختار کلی معماری پیشنهادی

در این پژوهش، تلاش شده است تا با توجه به دیدگاه‌های متنوع و روابط بین کلمات و جملات در یک سند، جملاتی با ارزش معنایی بالاتر استخراج شوند. برای این منظور، بخشی از سیستم به استخراج بردار ویژگی‌ها برای جملات یک سند می‌پردازد. ابتدا جملات و کلمات سند استخراج و پیش‌پردازش می‌شوند. برای تبدیل کلمات به بردارهای معنایی، از جانشانی‌های معروفی مانند Glove یا BERT استفاده می‌شود. در این مرحله، با توجه به اطلاعات زمینه‌ای، نحوی، معنایی و روابط بین موجودیت‌های متن در انواع سطح، تلاش برای استخراج مناسب اطلاعات انجام می‌شود و برای هر جمله بردار ویژگی مناسبی به دست می‌آید. در مرحله بعد، بردارهای جملات با استفاده از یک روش امتیازدهی مرتب شده تا رتبه‌ی معنایی هر جمله به دست آید. در نهایت، با استفاده از یک الگوریتم انتخاب جمله، جملات خلاصه استخراج می‌شوند. شکل ۳۰ نمایانگر روند کلی سیستم خلاصه‌سازی و استخراج متن است.



شکل ۳۰: روند کلی معماری خلاصه‌سازی خودکار متن با رویکرد استخراجی.

در این پژوهش، از تنوع روابط و شبکه‌های عصبی برای استخراج بردارهای ویژگی بهره‌مند شده است. علاوه بر این، برای ارزیابی جملات از یک شبکه کاملاً متصل استفاده شده، به نحوی که چند جمله ابتدایی با امتیاز بالاتر انتخاب و استخراج می‌شوند. استخراج بردار ویژگی مناسب برای جملات که تمام ویژگی‌های مناسب برای خلاصه‌سازی را داشته باشد، مهم‌ترین بخش در روند ذکر شده می‌باشد. در ادامه به بررسی ساختارهای ارائه شده در این پژوهش پرداخته می‌شود.

۴-۲- ساختار SWGraph^۱

در این پژوهش، از ساختار SWGraph با الهام از مدل HSG^۱، که در فصل ۳ معرفی شده است، برای استخراج ویژگی‌های در سطح سند در جملات استفاده شده است. این روش از یک گراف ناهمگن برای مدل‌سازی روابط بین جملات و کلمات در یک سند بهره برده است. در این گراف، جملات و کلمات به عنوان گره‌های گراف در نظر گرفته شده و بین هر جمله و کلمات موجود در آن جمله، یال قرار داده می‌شود. روند ساخت گراف SWGraph به شرح زیر است:

- ۱- استخراج جملات از متن ورودی.
- ۲- پیش‌پردازش جملات و حذف علائم نگارشی و کلمات توقف^۲.
- ۳- استخراج کلمات داخل هر جمله و ساخت لیستی از کلمات یکتا در متن.
- ۴- ساخت گراف با گره‌های جملات و کلمات یکتا و اتصال بدون جهت هر جمله به کلماتی که دارد.

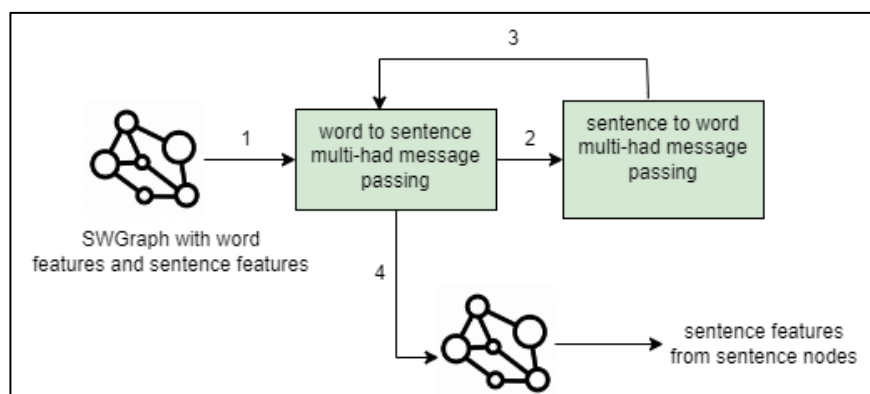
شکل ۳۱ نمونه‌ای از گراف تولید شده متناظر با متن داده شده را نشان می‌دهد.

پس از ایجاد گراف، باید ویژگی‌های گره‌های جملات و کلمات تعیین شود، برای بردار ویژگی کلمات از یک جانشانی کلمه و برای بردار ویژگی جملات از لایه‌های متنوع شبکه‌های عصبی و اطلاعات موجود در کلمات، استفاده می‌شود. همچنین به منظور بهبود تبادل اطلاعات و افزایش دقت یادگیری، از یک جانشانی قابل آموزش متناسب با ورودی مقدار TF-IDF کلمه و جمله، برای تولید بردار ویژگی یال‌ها استفاده می‌شود.

جهت آموزش گراف SWGraph، می‌توان از شبکه‌های عصبی گراف متنوع بهره گرفت. در این تحقیق، برای انتقال پیام بین گره‌ها از دو ساختار توجه چند سره استفاده شده است. در ابتدا، انتقال پیام از کلمات به جملات انجام می‌شود و سپس از جملات به کلمات، در نهایت دوباره از کلمات به جملات. این رویکرد باعث می‌شود تا در فرآیند خلاصه‌سازی، به کلمات با ارزش بیشتر اهمیت بیشتری تعلق گیرد. به علاوه، جملاتی که کلمات مشترک دارند، به واسطه کلمات مشترک، ویژگی‌های هم را به اشتراک می‌گذارند و بردارهای آن‌ها بیشتر به هم شبیه می‌شوند. با توجه به موارد ذکر شده، بردار ویژگی سطح سند برای جملات استخراج شده و در فرآیند انتخاب جمله در خلاصه‌سازی به کار می‌رود. شکل ۳۲ نمای کلی معماری SWGraph را نمایش می‌دهد.

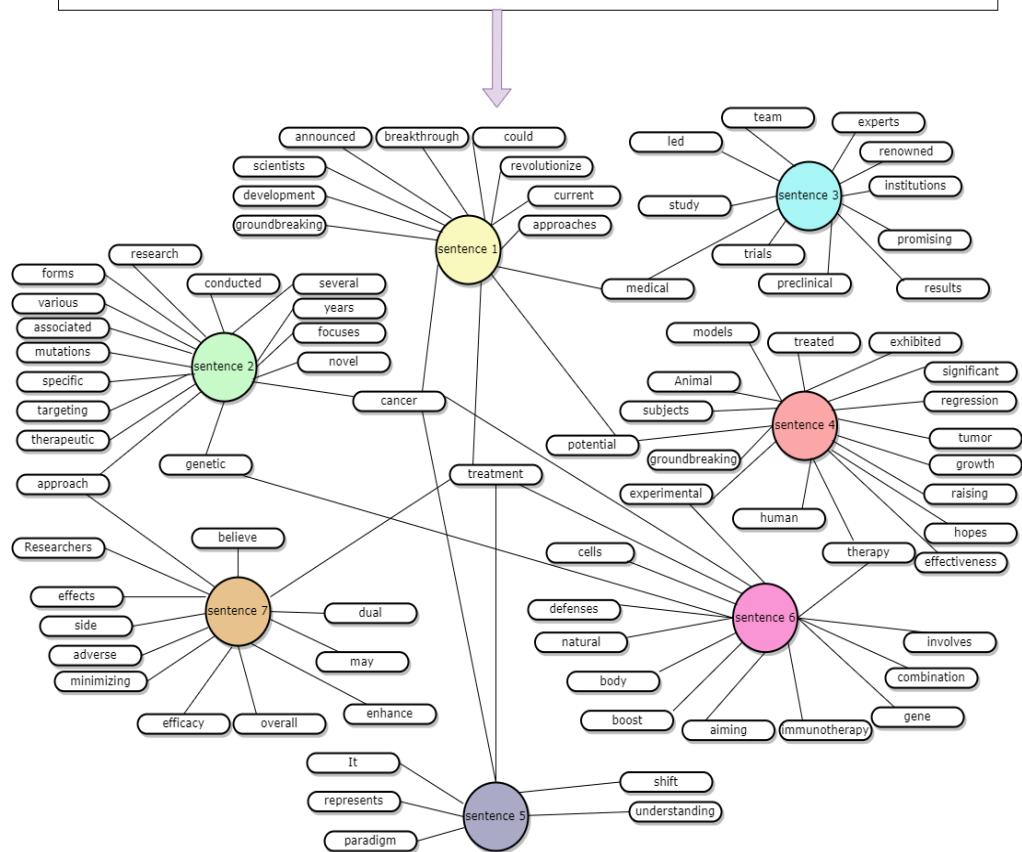
¹ Sentence-word graph

² Stop words



شکل ۳۱: نمای کلی معماری SWGraph برای بدست آوردن بردار ویژگی‌های جملات.

In a groundbreaking development, scientists have announced a potential breakthrough in cancer treatment that could revolutionize current medical approaches. The research, conducted over several years, focuses on a novel therapeutic approach targeting specific genetic mutations associated with various forms of cancer. The study, led by a team of experts from renowned medical institutions, demonstrates promising results in preclinical trials. Animal models treated with the experimental therapy exhibited significant regression in tumor growth, raising hopes for the potential effectiveness in human subjects. It represents a paradigm shift in our understanding of cancer and its treatment. The experimental treatment involves a combination of gene therapy and immunotherapy, aiming to boost the body's natural defenses against cancer cells. Researchers believe that this dual approach may enhance the overall efficacy of the treatment while minimizing adverse side effects.

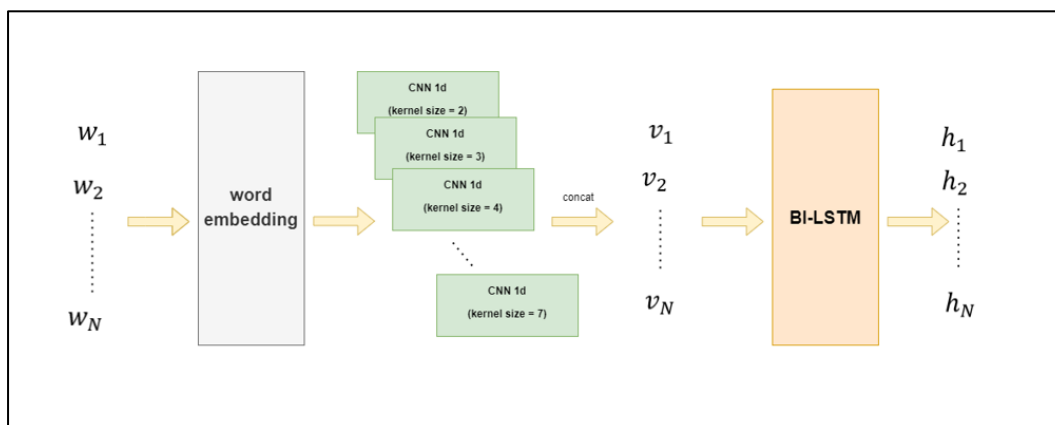


شکل ۳۳: نمونه‌ای گراف *SWGraph* برای یک متن.

۴-۳- ساختار CNN-LSTM

ساختار CNN-LSTM حاصل انواع ترکیب از دو مدل عصبی معروف، یعنی شبکه‌های عصبی CNN و شبکه‌های عصبی LSTM است که معمولاً در پردازش متن دنباله‌ای مانند جملات یا سندها به کار می‌رود. در این ساختار، لایه‌های CNN برای استخراج ویژگی‌ها از دنباله کلمات متن به کار می‌روند. این لایه‌ها با استفاده از فیلترها و عملیات پیچش، الگوها و ویژگی‌های مهم در جملات را شناسایی می‌کنند. به عبارت دیگر، CNN مسئول درک و تجزیه و تحلیل ویژگی‌های مکانی کلمات در جمله است. سپس، خروجی لایه‌های CNN به عنوان ورودی به لایه‌های LSTM ارسال می‌شود. لایه‌های LSTM از حافظه کوتاه‌مدت و حافظه بلندمدت برای نگهداری اطلاعات وابسته به زمان استفاده می‌کنند. این لایه‌ها می‌توانند روابط و ویژگی‌های دنباله‌ای را در جمله مدیریت کرده و درک عمیق‌تری از ساختار جمله را ارائه دهند. با ترکیب این دو مدل می‌توان از ویژگی‌های مکانی متن به دست آمده توسط CNN به همراه توانایی LSTM در درک وابستگی‌های زمانی متن، برای وظایفی نظیر خلاصه‌سازی متون یا تشخیص اجزاء مهم در متون بهره‌مند شد. این ساختار معمولاً در وظایف پردازش متن و تولید خروجی‌های دنباله‌ای با توجه به ساختار و معنا در متون، کاربرد دارد [30].

همانطور که در شکل ۳۳ آمده است، مجموعه‌ای از لایه‌های پیچشی یک بعدی با سایز پنجره‌های مختلف به منظور شبیه‌سازی ویژگی‌های n-gram و استخراج ویژگی‌های مهم از دنباله کلمات، بر روی بردارهای کلمات یک جمله اعمال می‌شوند. سپس با استفاده از لایه Bi-LSTM، ویژگی‌های دنباله‌ای و زمینه‌ای جمله استخراج می‌شود.

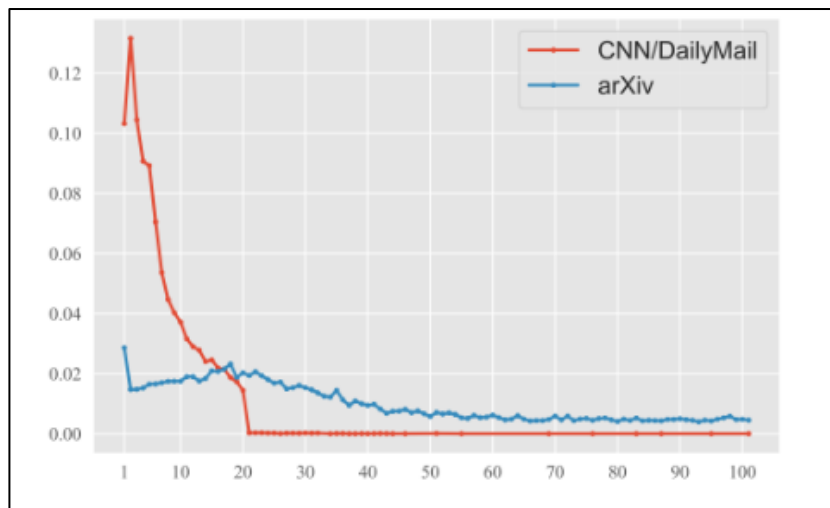


شکل ۳۳: شمای کلی ساختار CNN-LSTM در مدل پیشنهادی.

۴-۴- ساختار جانشانی مکانی^۱

در یک ساختار منسجم انگلیسی، موقعیت یک جمله در متن قادر است تا میزان اهمیت آن را به وضوح اظهار کند. همانگونه که در نمودار ۳۴ آمده است، جملات ابتدایی معمولاً از اهمیت بیشتری برخوردارند. به علاوه، جملاتی که در انتهای متن قرار می‌گیرند، معمولاً بار معنایی اصلی متن را حمل نمی‌کنند. به همین منظور در این پژوهش ساختار SentPosEmbed معرفی شده است که برای جانشانی مکانی جمله در متن را به عنوان یک ویژگی در نظر می‌گیرد.

جانشانی مکانی، یک مفهوم مهم در زمینه یادگیری عمیق است که در مدل‌هایی مانند شبکه‌های ترنسفر به کار می‌رود. جانشانی مکانی به ما کمک می‌کند تا اطلاعات مکانی یا ترتیبی جمله یا دنباله‌ای از واحدهای ورودی را به مدل ارائه دهیم. وقتی با داده‌های توالی، مانند جملات یا دنباله‌های زمانی، سر و کار داریم مکان یا ترتیب کلمات یا جملات اطلاعات مهمی ارائه می‌دهد [43].



شکل ۳۴: میزان وقوع حضور جملات در خلاصه با جایگاه مکانی متفاوت در سند داخل دو مجموعه داده CNN/DailyMail و arXiv [75].

در ساختار مدل‌های ترنسفرمر، اطلاعات ترتیبی یا مکانی به منظور افزودن به واحدهای ورودی، از توابع سینوسی و کسینوسی بهره می‌برند. با در نظر گرفتن ورودی، یک جدول کدگذاری سینوسی^۲ برای هر مکان ممکن در دنباله ایجاد می‌شود. این جدول سپس به موقعیت مکانی در دنباله متناظر می‌شود و اطلاعات مکانی به بردار ویژگی تبدیل می‌گردد. این رویکرد، به مدل این امکان را می‌دهد

¹ Position embedding

² sinusoid_encoding_table

که ترتیب و مکان اجزای ورودی را بهبود بخشیده و اطلاعات ترتیبی معنی‌دارتری را ادغام کند. روابط ریاضی ۲۳ و ۲۴، نحوه‌ی محاسبات جانشانی موقعیت را در ساختارهای ترنسفرمر نشان می‌دهد.

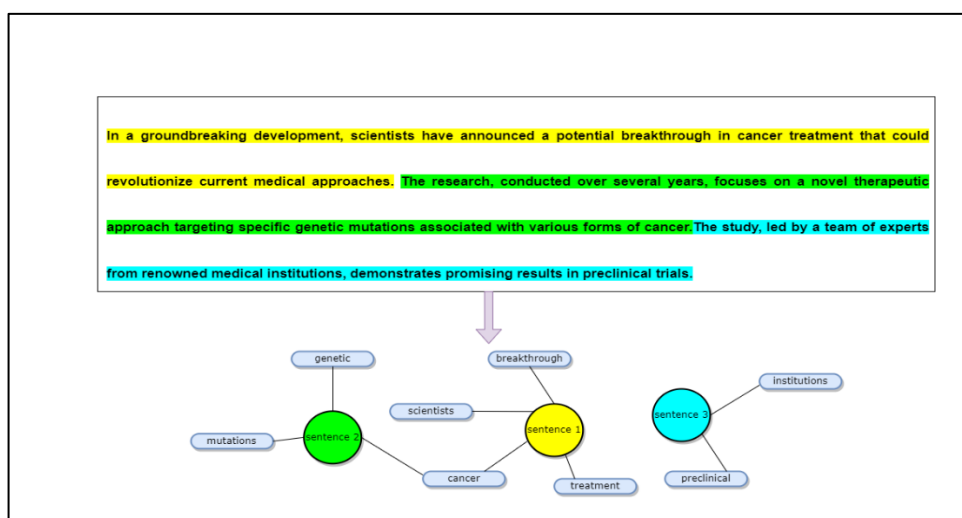
$$PositionEmbed(pos, 2i) = \sin\left(\frac{pos}{1000^{\frac{2i}{d}}}\right) \quad (23)$$

$$PositionEmbed(pos, 2i + 1) = \cos\left(\frac{pos}{1000^{\frac{2i}{d}}}\right) \quad (24)$$

در معادله‌های بالا، pos نشان‌دهنده‌ی موقعیت مکانی جمله، i نمایان‌گر اندیس بُعد ورودی، d بُعد کل می‌باشد. عددی که از طریق یک مدل پیش‌آموزش دیده شده برای جانشانی موقعیت به دست می‌آید، به یک بردار تبدیل می‌شود و در فرآیند یادگیری این مدل جانشانی موقعیت، مقادیر مرتبط با مکان آن آموزش می‌بیند [76].

۴-۵- ساختار KeywordGraph

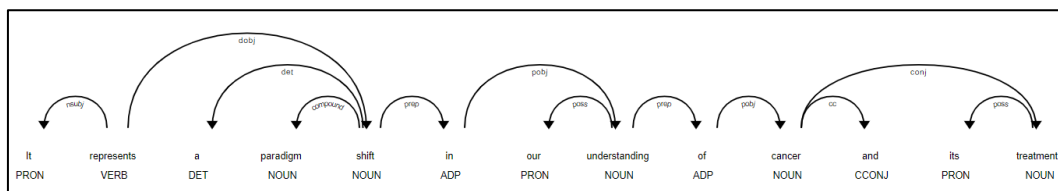
همانطور که در فصل ۳ توضیح داده شد، یکی از رایج‌ترین روش‌های موجود در خلاصه‌سازی خودکار متن، استفاده از عناوین یا کلیدواژه‌ها می‌باشد. در این مدل‌ها، استخراج کلیدواژه‌ها یا عناوین بسیار اهمیت دارد و در برخی مواقع اشتباه در استخراج این موارد باعث می‌شود دقت مدل پایین آید. همچنین چگونگی استفاده از این عناوین یا کلیدواژه‌ها برای خلاصه‌سازی نیز اهمیت بسیار زیادی دارد. در این پژوهش از ابزار keyBert به منظور استخراج کلیدواژه‌ها استفاده شده است. به مشابه SWGraph گرافی شامل کلیدواژه‌ها و جملات تهیه می‌شود که نشان دهنده‌ی آن است که چه جملاتی حاوی چه کلیدواژه‌هایی هستند. با استفاده از شبکه‌ی عصبی گرافی مبتنی بر توجه، انتقال پیام صورت گرفته و بردار ویژگی‌های شبکه استخراج می‌شود. شکل ۳۵ نمونه‌ای از کلیدواژه‌های استخراج شده و نحوه‌ی ساخت گراف را نشان می‌دهد.



شکل ۳۵: کلیدواژه‌های استخراجی از یک نمونه متن و ساخت گراف متناظر با آن.

به منظور بهبود درک بهتر جملات توسط الگوریتم، استفاده از اطلاعات نحوی مفید واقع شوند. در فرآیند خلاصه‌سازی متن، ارتباطات نحوی بین کلمات و جملات نقش اساسی ایفا می‌کنند. این اطلاعات نحوی به مدل‌های خودکار خلاصه‌سازی کمک می‌کنند تا ساختار جملات و ارتباطات میان کلمات را بهتر درک کرده و جملات حاوی اطلاعات حیاتی را به عنوان جملات کلیدی برجسته کنند. به عنوان مثال، درخت نحوی یک جمله را به صورت یک ساختار درختی، ارتباطات نحوی بین کلمات و جملات را نمایش می‌دهد. این اطلاعات می‌توانند برای تشخیص ویژگی‌های مهم جملات و اهمیت آنها در متن بهره‌ور باشند. عناصری نظیر وابستگی‌ها، نقش‌های نحوی و ساختار جمله می‌توانند به مدل‌های خلاصه‌سازی کمک کنند تا جملات حاوی اطلاعات مهم و تأثیرگذار در متن را انتخاب کرده و آنها را به صورت خلاصه به کار ببرند. همچنین، دانش از ساختار نحوی می‌تواند اهمیت موجودیت‌های مختلف را بیشتر مورد بررسی قرار داده و به کلماتی که از نظر نحوی اهمیت بیشتری دارند، توجه بیشتری اختصاص دهد.

در این پژوهش، با الهام از یک مدل در مسئله‌ی تشخیص احساسات^۱ [77]، جهت بهره‌وری از اطلاعات نحوی، از ابزار Spacy بهره گرفته شده است. Spacy یک کتابخانه پردازش زبان طبیعی است که امکانات گسترده‌ای برای تجزیه و تحلیل متون، از جمله قابلیت ساخت درخت نحوی، ارائه می‌دهد. شکل ۳۶ یک نمونه از درخت خروجی از ابزار Spacy می‌باشد.



شکل ۳۶: مثالی از گراف نحوی تولید شده برای یک جمله.

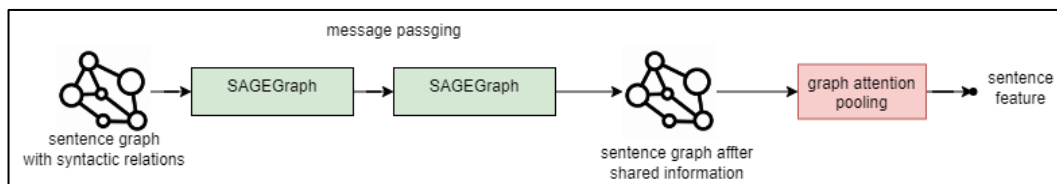
پس از ساخت گراف نحوی برای هر جمله، از یک شبکه‌ی عصبی گرافی برای اشتراک ویژگی‌ها استفاده شده و بعد از آن از یک روش ادغام نیز به منظور استخراج ویژگی هر جمله استفاده می‌شود. ساختار کلی SentSyntaxGraph به صورت شکل ۳۷ آمده است. در این ساختار از دو لایه‌ی SageGraph به منظور یادگیری پیکره‌بندی گراف و اشتراک اطلاعات گره‌ها استفاده شده است. همچنین از یک لایه‌ی graph-attention-pooling به منظور ادغام گره‌های گراف جمله و استخراج بردار ویژگی با استفاده از کلمات موجود در گراف، استفاده شده است. این روش ادغام سازی، مبتنی بر توجه است و در فرایند یادگیری میزان اهمیت کلمات مهم و اطلاعات مهم پیکره‌بندی یاد گرفته می‌شود.

¹ Sentiment analysis

۴-۷- ساختار SentSemanticGraph

در متون، ارتباطات معنایی^۱ نقش بنیادی در تداخل و اتصال مفاهیم و اطلاعات دارند. در فرآیند خلاصه‌سازی خودکار، توجه به این ارتباطات معنایی می‌تواند به مدل‌ها کمک کند تا بخش‌های اساسی و مؤثر متن را با دقت شناسایی نمایند. با درک و بهره‌گیری از این ارتباطات معنایی، مدل‌های خودکار قادر خواهند بود جملات یا بخش‌هایی از متن را انتخاب کنند که نقش بنیادی در انتقال مفهوم و اطلاعات اصلی دارند. این نکته به بهبود کیفیت و محتوای خلاصه‌های تولیدی ارتقا می‌بخشد. ساختار گراف معنایی ارتباطات معنایی میان موجودیت‌های یک متن را نمایان می‌کند. در خلاصه‌سازی خودکار، استفاده از گراف معنایی می‌تواند به مدل‌ها کمک کند تا ارتباطات معنایی میان مفاهیم را در نظر بگیرند و در نتیجه، بهترین تصمیمات را برای انتخاب و ترکیب اطلاعات کلیدی در خلاصه‌ها بگیرند. این رویکرد باعث بهبود یکنواختی و ارتباط مناسب میان عناصر مختلف متن می‌شود، که در نتیجه موجب تولید خلاصه‌هایی با ساختار منطقی و مفهومی تر می‌گردد.

در این پژوهش، با الهام از یک مدل ارائه شده در مسئله‌ی تشخیص احساسات در متن [77]، جهت تبدیل جملات به یک بردار ویژگی از یک گراف معنایی استفاده می‌شود. ساختار کلی این بخش، مشابه به ساختار SentSyntaxGraph است که در شکل ۳۷ نمایش داده شده است. برای ایجاد گراف معنایی، ابتدا کلمات جملات با استفاده از یک جانشانی کلمه به بردارهای ویژگی تبدیل و سپس، بین هر دو کلمه، اگر شباهت کسینوسی بیشتر از یک حد آستانه (به عنوان مثال ۰/۷) باشد، یک یال بین آن دو کلمه ایجاد می‌شود. این فرآیند برای کلمات مشابه در یک جمله تکرار شده و گراف معنایی برای جمله ساخته می‌شود. سپس، با استفاده از لایه‌های عصبی گراف، مانند SAGEGraphs، اطلاعات بین کلمات یک جمله اشتراک گذاری و در نهایت، با استفاده از لایه‌ی ادغام مبتنی بر توجه، بردار ویژگی از گراف جمله استخراج می‌شود.



شکل ۳۷: معماری ساختار SentSyntaxGraph و SentSemanticGraph

¹ Semantic relations

۴-۸- ساختار DocSemanticGraph

استفاده از بردار ویژگی معنایی برای کل سند، به عنوان یک نماینده کلان از محتوای متن، به مزایای مهمی در فرآیند خلاصه سازی خودکار متن منجر می شود. با استفاده از یک بردار ویژگی معنایی برای کل سند، مدل های خودکار قادرند تا اطلاعات مهم و اساسی در متن را به طور کلان شناسایی کنند. این کار باعث می شود تا خلاصه تولیدی، تمرکز بیشتری بر روی نقاط کلیدی و مفهومی تر متن داشته باشد، و اطلاعات اصلی کمتر گمراه شوند. این بهبود کیفیت و دقت خلاصه ها را افزایش می دهد. همچنین با تبدیل کل سند به یک بردار ویژگی معنایی، ابهامات معنایی کاهش می یابد و مدل بهتر می تواند ارتباطات معنایی بین اجزای مختلف متن را درک کند. این امکان را می دهد تا در فرآیند خلاصه سازی، مدل با دقت بیشتری اطلاعات مهم را انتخاب کرده و خلاصه هایی با ساختار منظم تر و فهم پذیرتر تولید کند. از این رو، افزایش فهم و انتقال معانی واقعی متن در خلاصه ها حاصل می شود.

در این پژوهش، به منظور در نظر گرفتن بردار ویژگی معنایی برای کل سند، یک گراف معنایی جامع بر روی کلمات منحصر به فرد متن ایجاد می شود. روش ساخت این گراف تطابق دارد با بخش پیشین و از شباهت کسینوسی بین کلمات متن برای ایجاد ارتباطات معنایی استفاده می شود. به علاوه، همانند بخش قبل، گره های کلمات با یکدیگر اطلاعات را به اشتراک می گذارند و سپس با استفاده از لایه ای ادغام مبتنی بر توجه، برای کل سند یک بردار ویژگی استخراج می شود. این بردار ویژگی برای تمام جملات یک سند یکسان است و به بردار ویژگی جملاتی که با روش های دیگر استخراج شده اند، متصل می شود و وارد لایه ای امتیازدهی می شود. این رویکرد منجر به ایجاد یک نمایش یکپارچه و با کیفیت از معنای کل سند می گردد.

۴-۹- مدل های ارائه شده

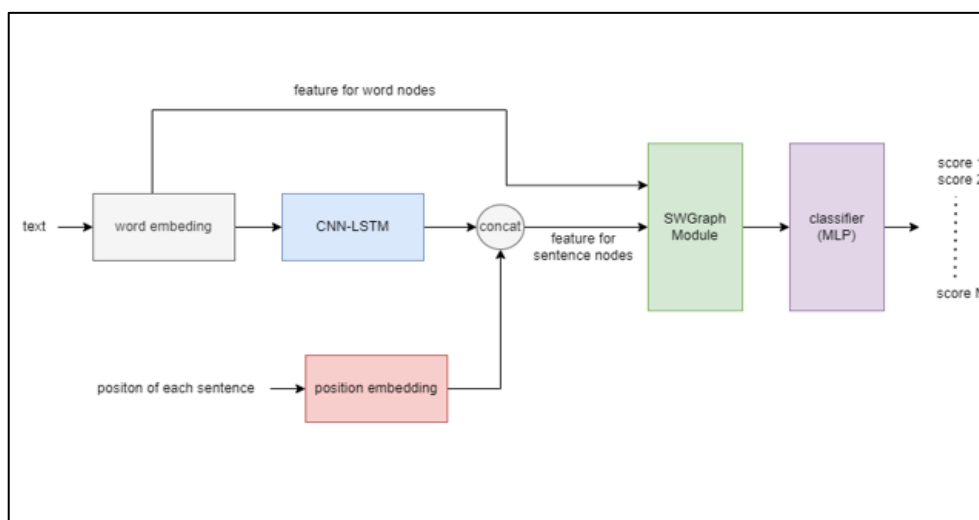
در بخش های قبل، در مورد معماری کلی مدل های پیشنهادی و بخش های مختلف آن پرداخته شد. هدف از این بخش آن است تا مدل ها پیشنهادی نهایی و چگونگی استفاده از لایه های ذکر شده توضیح داده شود.

همان طور که پیش تر توضیح داده شد، ابتدا متن را به عناصر قابل پردازش تقسیم می کنیم. متن ورودی به جملات آن و هر جمله به کلمات آن شکسته می شود. با استفاده از روش برچسب گذاری Oracle که پیش تر توضیح داده شد، جملات استخراج شده مطابق با خلاصه ی موجود، به

برچسب‌های صفر و یک منطبق شده‌اند. این اقدام باعث تبدیل مسئله به یک مسئله‌ی دسته‌بندی شده است.

با استفاده از لایه‌های متعددی که در بخش‌های پیشین ارائه شدند، برای هر جمله بردارهای ویژگی تولید می‌شوند. سپس، با به کارگیری از لایه‌ی امتیازدهی، که یک شبکه‌ی عصبی MLP است، جملات امتیازدهی و مرتب می‌گردند. در این فرآیند با استفاده از تابع خطای CrossEntropy مقادیر خطا برای هر جمله محاسبه و آموزش داده می‌شوند. در مرحله ارزیابی، جملات با امتیاز بالاتر استخراج شده (برای مثال ۴ جمله با امتیاز بالاتر) و متن خلاصه تولید می‌گردد.

شکل ۳۸ معماری اولیه و پایه ارائه شده را به نمایش می‌کشد. در این معماری از یک گراف SWGraph به منظور ساخت بردار ویژگی جملات استفاده شده و خلاصه تولید می‌شود. بردار ویژگی‌های ورودی گره‌های جملات از اتصال ویژگی‌های CNN-LSTM کلمات داخل جمله با ساختار جانشانی مکانی که در بخش‌های قبل توضیح داده شده، تشکیل شده است. این معماری، ویژگی‌های سطح سند را استخراج می‌کند. همچنین با توجه به استفاده از ساختار توجه چند سره در یادگیری گراف SWGraph، کلمات مهم‌تر در سطح سند، در تصمیم‌گیری اهمیت بیشتری خواهند داشت.



شکل ۳۸: مدل اولیه با استفاده از ساختار SWGraph.

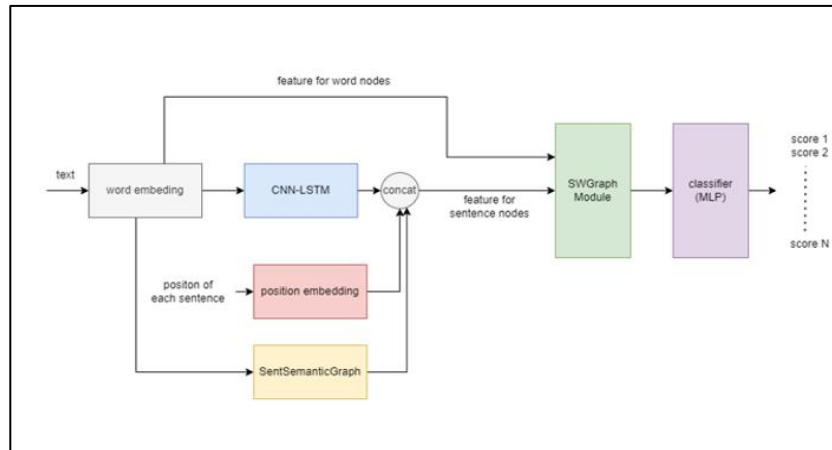
در کنار مدل SWGraph، برای بهبود استفاده از ویژگی‌های سطح جمله، از گراف‌های مبتنی بر ساختار جمله مانند SentSyntaxGraph و SentSemanticGraph می‌توان بهره گرفت. این گراف‌ها قابلیت استخراج ویژگی‌های موثر را بر اساس روابط تعریف شده بین کلمات در جمله و نیز ساختارهای زبانی مرتبط با آن‌ها دارند. این اطلاعات مفید می‌توانند به تمایز و استخراج ویژگی‌های مهم‌تر هر جمله کمک کرده و بهبود قابل توجهی در عملکرد مدل ایجاد کنند.

همچنین به منظور توجه بیشتر به کلیدواژه‌ها و عناوین می‌توان از ساختارهای KeywordGraph یا DocSemanticGraph استفاده می‌شود. KeywordGraph باعث می‌شود تا کلیدواژه‌ها به صورت واضح به مدل شناسانده شوند. ساختار DocSemanticGraph نیز می‌توان بردارویژگی در سطح سند با توجه به کلمات مهم‌تر ایجاد کند.

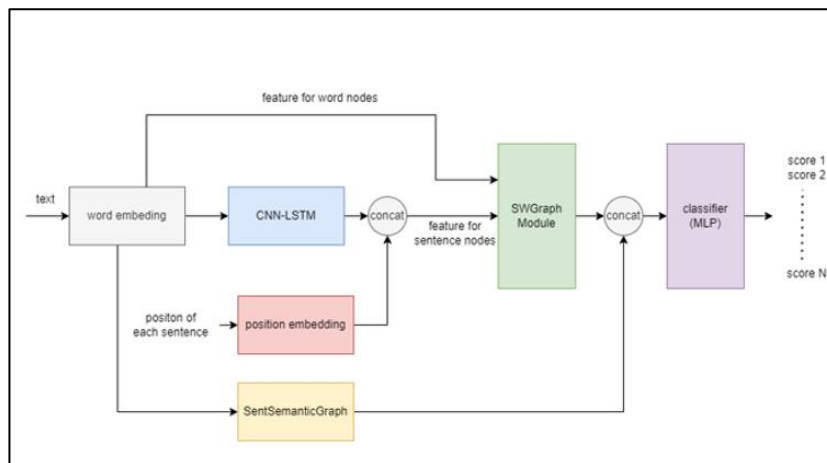
تمام ساختارها و ترکیبات ذکر شده می‌توانند به عنوان استخراج کننده‌ی بردارهای ویژگی جملات استفاده شود. به همین منظور انواع معماری‌ها آموزش و آزمایش شدند. نتایج مدل‌ها در فصل ۵ آمده است، که بهترین آن‌ها به شرح زیر است:

- ۱- SWGraph + SentSemanticGraph به صورت توالی
- ۲- SWGraph + SentSemanticGraph به صورت موازی
- ۳- SWGraph + SentSyntaxGraph
- ۴- SWGraph + KeywordGraph
- ۵- SWGraph + DocSemanticGraph
- ۶- SWGraph + SentSemanticGraph + SentSyntaxGraph
- ۷- SWGraph + SentSemanticGraph + KeywordGraph
- ۸- SWGraph + SentSemanticGraph + DocSemanticGraph

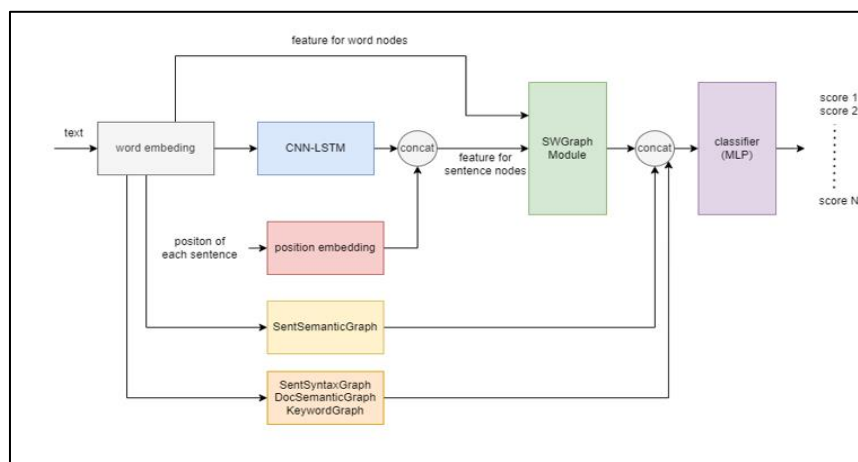
در این بین بهترین نتیجه برای ترکیب مدل SWGraph با SentSemanticGraph بوده است که نتایج آن در فصل ۵ آمده است. می‌توان خروجی ساختار SentSemanticGraph را با استفاده از خروجی ساختارهای CNN-LSTM و Position-Embedding متصل نمود و به عنوان بردار ویژگی ورودی ساختار SWGraph استفاده نمود (شکل ۳۹). همچنین می‌توان خروجی ساختار SentSemanticGraph را با خروجی ساختار SWGraph متصل نمود و به عنوان بردار ویژگی جمله به دسته‌بند داد (شکل ۴۰). ساختار موازی SentSemanticGraph و SWGraph بهترین نتیجه را در بین آزمایش‌ها داشته و به عنوان مدل نهایی این پژوهش معرفی می‌شود. در ادامه نیز تمام ترکیبات ساختارهای SWGraph و SentSemanticGraph با بقیه ساختارها نیز آزمایش و گزارش شده‌اند (شکل ۴۱).



شکل ۳۹: ترکیب ساختار *SentSemanticGrph* با *SWGraph* به صورت توالی



شکل ۴۰: ترکیب ساختار *SentSemanticGrph* با *SWGraph* به صورت موازی



شکل ۴۱: ترکیب مدل *SWGraph* و *SentSemanticGraph* با دیگر مدل‌های استخراج‌کننده‌ی بردار ویژگی جملات.

۴-۱۰- جمع‌بندی

در این بخش، روش‌های پیشنهادی خلاصه‌سازی خودکار متن با تأکید بر رویکرد استخراجی مورد بررسی قرار گرفت. مفاهیم اساسی برای پیاده‌سازی سیستم خلاصه‌سازی ابتدا معرفی شدند و سپس مسئله خلاصه‌سازی متن به عنوان یک مسئله دسته‌بندی دوتایی با رویکرد استخراجی تعریف شد. ارتباطات بین عناصر مختلف به عنوان ویژگی‌های کلیدی مورد توجه قرار گرفت و اطلاعات زمینه‌ای، اطلاعات معنایی و نحوی نیز در سطوح مختلف برای ساخت خلاصه مطرح شدند.

در معماری پیشنهادی این پژوهش، تمرکز بر استخراج جملات با ارزش معنایی بالا از سند با در نظر گرفتن دیدگاه‌های متنوع و روابط میان کلمات و جملات قرار دارد. در مرحله پیش‌پردازش، از جانشانی‌های کلمه معروف مانند Glove یا BERT برای تبدیل کلمات به بردارهای معنایی استفاده می‌شود. با توجه به اطلاعات زمینه‌ای، نحوی، معنایی و روابط بین موجودیت‌های متن در انواع سطح، بردارهای ویژگی برای هر جمله استخراج می‌شوند. بردارهای جملات با استفاده از یک روش امتیازدهی مرتب شده و جملات با بیشترین اهمیت معنایی برای استفاده در خلاصه‌سازی انتخاب می‌گردند.

در این پژوهش از ساختارهای متنوعی به جهت بهبود خلاصه‌سازی خودکار متن استفاده شده است. ساختار SWGraph در این پژوهش از یک گراف ناهمگن الهام‌گرفته از مدل HSG برای مدل‌سازی روابط بین جملات و کلمات در متن استفاده می‌کند. در این گراف، جملات و کلمات به عنوان گره‌ها در نظر گرفته شده و ارتباطات میان آنها با یال‌های گراف نمایان می‌شود. سپس با استفاده از جانشانی‌های معروف برای کلمات و لایه‌های متنوع شبکه‌های عصبی برای جملات، ویژگی‌های گره‌ها تعیین می‌شود. این ساختار با بهره‌گیری از جانشانی‌های قابل آموزش و الهام‌گیری از مدل TF-IDF برای بهبود تبادل اطلاعات، به ساخت یک نمایش یکپارچه از متن با تمرکز بر روابط معنایی بین اجزای آن می‌پردازد.

ساختار CNN-LSTM این پژوهش، از ترکیب مدل‌های عصبی CNN و LSTM برای پردازش متون دنباله‌ای استفاده می‌کند. لایه‌های CNN برای استخراج ویژگی‌های مکانی کلمات به کار می‌روند و سپس این ویژگی‌ها به لایه‌های LSTM منتقل می‌شوند. این ساختار ترکیبی باعث بهبود درک اطلاعات مکانی و زمانی متن می‌شود و برای وظایف مانند خلاصه‌سازی بسیار موثر است.

مدل‌های ارائه شده، از جدول کدگذاری سینوسی به عنوان جانشانی مکانی جملات برای بهبود درک ترتیب و مکان جملات در متن استفاده می‌کنند. این ابزار با انتقال اطلاعات مکانی به مدل، ترتیب و موقعیت اجزای متن را بهتر درک می‌کند و در وظایفی که ترتیب اطلاعات مهم است، نقش بسیار موثری ایفا می‌کند.

در روش‌های ارائه شده، از اطلاعات نحوی به منظور بهبود درک جملات استفاده می‌شود. با استفاده از ابزار Spacy، اطلاعات نحوی از جملات استخراج شده و به شکل گراف نحوی نمایان می‌شود. سپس از شبکه‌های عصبی گرافی برای اشتراک ویژگی‌ها و استخراج ویژگی هر جمله استفاده می‌شود. این ساختار با دقت بالا ارتباطات و ساختار نحوی جملات را در مدل‌های خودکار خلاصه‌سازی در نظر می‌گیرد.

در این پژوهش با استفاده از گراف معنایی، ارتباطات معنایی بین موجودیت‌های یک متن نمایش داده می‌شود. از جانشانی کلمه برای تبدیل کلمات به بردارهای ویژگی استفاده شده و با ایجاد گراف معنایی بر مبنای شباهت کسینوسی، ارتباطات معنایی بین کلمات مدل می‌شود. این ساختار با بهبود درک معنایی جملات و اطلاعات مهم متن، در مدل‌های خلاصه‌سازی بهره‌ور استفاده می‌شود.

در ساختار keywordGraph، از اهمیت و استخراج کلیدواژه‌ها برای خلاصه‌سازی متن استفاده می‌شود. از ابزار keyBert جهت استخراج کلیدواژه‌ها استفاده شده و سپس با ایجاد یک گراف که شامل کلیدواژه‌ها و جملات متن است، ارتباطات بین آنها نمایان می‌شود. از یک شبکه‌ی عصبی گرافی برای انتقال پیام و استخراج بردارهای ویژگی استفاده می‌شود.

معماری DocSemanticGraph نیز، جهت بهبود درک معنایی کل سند توسط الگوریتم خودکار خلاصه‌سازی استفاده می‌شود. با ایجاد گراف معنایی جامع بر روی کلمات منحصر به فرد متن، ارتباطات معنایی بین آنها نمایش داده می‌شود. این گراف معنایی با استفاده از لایه‌های ادغام مبتنی بر توجه و بردارهای ویژگی معنایی، به مدل‌ها کمک می‌کند تا اطلاعات مهم و اساسی در متن را به‌طور کلان‌شناسایی کنند و در فرآیند خلاصه‌سازی بر اساس این اطلاعات عمل کنند.

در این پژوهش، مدل‌های خودکار خلاصه‌سازی با هدف بهبود درک جملات و ارتباطات معنایی و نحوی متن توسعه یافته‌اند. از ساختارهای متنوع ذکر شده به همراه ترکیبات مختلف این ساختارها برای استخراج ویژگی‌های جملات و تولید خلاصه‌های متن بهره گرفته شده است. ابتدا متن به عناصر قابل پردازش تقسیم می‌شود و سپس جملات با استفاده از برچسب‌گذاری Oracle به برچسب‌های صفر و یک بر اساس خلاصه‌ی موجود در متن تبدیل می‌شوند. برای هر جمله، از لایه‌های ذکر شده استفاده و بردارهای ویژگی تولید می‌شوند. در مرحله آموزش مدل‌ها، از لایه‌های امتیازدهی و تابع خطای CrossEntropy استفاده شده و در فرآیند ارزیابی، جملات با امتیاز بالاتر جهت تولید خلاصه انتخاب می‌شوند. نتایج به دست آمده از ترکیب‌های مختلف مدل‌ها گزارش شده و بهترین عملکرد برای ترکیب SWGraph با SentSemanticGraph حاصل شده است. نتایج بدست آمده از مدل‌های ارائه شده در فصل بعد گزارش شده است. این تحقیق با تأکید بر اطلاعات نحوی و معنایی، بهبود یکنواختی و ارتباط مناسب بین عناصر مختلف متن را هدف اصلی خود قرار داده است و توانسته خلاصه‌های بهتری تولید کند.

فصل پنجم

آزمایش و نتایج

۵-۱- مقدمه

در این فصل، به تفصیل به بررسی جزئیات آزمایش‌ها و نتایج به‌دست‌آمده از پژوهش حاضر پرداخته خواهد شد. هدف اساسی این آزمایشات ارزیابی کارایی مدل‌های خودکار خلاصه‌سازی پیشنهادی بوده و نتایج به دست آمده با سایر روش‌ها جهت مقایسه مورد بررسی قرار گرفته‌اند. این بررسی شامل جنبه‌های گوناگون مدل‌ها، همراه با تحلیل عملکرد آن‌ها می‌شود. همچنین، در این بخش به جزئیات پیاده‌سازی و ابزارهای مورد استفاده در این پژوهش پرداخته خواهد شد و مجموعه داده مورد استفاده نیز معرفی خواهد شد. در ادامه، فرآیندهای پیش‌پردازش، تنظیم پارامترها، مراحل یادگیری، و تحلیل جزئی نتایج به دقت مورد بررسی قرار گرفته و نتایج به‌دست‌آمده با سایر مدل‌ها جهت ارزیابی و مقایسه دقیق‌تر بررسی خواهند شد.

این تحلیل گسترده نه تنها شامل مقایسه نتایج کمی ارائه شده توسط مدل‌هاست، بلکه به تبیین چگونگی عملکرد هر مدل از نظر کیفی نیز پرداخته و این امکان را فراهم می‌کند که از جوانب مختلف پژوهش بهره‌مند شویم. همچنین، این فصل به تشریح روند انجام آزمایشات، مشکلات موجود در هنگام انجام تحقیقات و راه‌حل‌های اتخاذ شده برای آنها خواهد پرداخت.

۵-۲- جزئیات پیاده‌سازی

در این پژوهش، برای پیاده‌سازی از زبان برنامه‌نویسی Python و از چارچوب^۱ Pytorch به منظور آموزش شبکه‌های عصبی استفاده شده است. از ابزار DGL برای ساخت و آموزش گراف‌های بزرگ، از ابزار nltk برای پیش‌پردازش داده‌ها، از ابزار Spacy برای ساخت گراف نحوی، و از ابزار keyBert برای استخراج کلیدواژه‌های متون استفاده گردیده است. همچنین، به منظور جانشانی کلمات از جانشانی‌های معروف مانند Glove و BERT بهره گرفته شده است.

برای بهبود سرعت آموزش، از واحد پردازش گرافیکی^۲ (GPU) با مشخصات مدل ۳۰۷۰ و رم گرافیکی ۸ گیگابایت استفاده گردیده است. همچنین، گراف‌های مورد نظر پیش‌تر ساخته شده و در قالب فایل‌های مناسب ذخیره گردیدند، که این اقدام باعث افزایش چشمگیر در سرعت آموزش مدل گردیده است. این ترکیب از ابزارها و فناوری‌ها، نه تنها به بهبود کارایی مدل‌ها کمک نموده است بلکه به شدت به کارایی و سرعت فرآیند آموزش و تحلیل داده‌ها افزوده است.

۵-۳- مجموعه داده

به منظور آموزش و ارزیابی مدل‌های پیشنهادی برای مسئله‌ی خلاصه‌سازی متن از مجموعه دادگان CNN/Daily Mail استفاده شده است. مجموعه دادگان CNN/Daily Mail یک مجموعه داده معروف در زمینه خلاصه‌سازی متن است که شامل اخبار و مقالات خبری از سایت‌های خبری CNN و DailyMail می‌باشد. این مجموعه داده شامل متون مرتبط با اخبار و رویدادهای روز به همراه خلاصه‌های مرتبط با هر متن است. این مجموعه داده به عنوان یک مرجع ارزیابی^۳ مهم در زمینه‌های پردازش زبان طبیعی و خلاصه‌سازی مورد استفاده قرار گرفته و ارزیابی مدل‌ها و روش‌های مختلف در این حوزه را تسهیل می‌کند. در جدول ۲ تعداد سند در هر بخش از هر مجموعه داده‌ی CNN و Daily Mail ارائه شده است. همچنین میانگین تعداد کلمه در هر سند نیز در جدول ۳ ذکر شده است. جدول ۴ نیز میانگین تعداد جملات در متن ورودی و میانگین تعداد جملات منتخب خلاصه در بخش‌های مختلف نشان می‌دهد [11].

¹ Framework

² Graphic process unit

³ benchmark

جدول ۲: تعداد سندها در مجموعه داده CNN/DailyMail [11].

	CNN	DAILYMAIL	مجموع
آموزش	90,266	196,961	287,227
ارزیابی	1,220	12,148	13,368
آزمون	1,093	10,397	11,490

جدول ۳: میانگین تعداد کلمات در هر سند [11].

	CNN	DAILY MAIL
آموزش	762	813
ارزیابی	763	774
آزمون	716	780

جدول ۴: میانگین تعداد جملات در هر سند.

	سند ورودی	خلاصه
آموزش	32	3.4
ارزیابی	35	3.5
آزمون	36	3.4

۴-۵- پیش پردازش

برای آماده سازی متون برای وارد شدن به شبکه های عصبی، نیاز به انجام یک سری عملیات پیش پردازش وجود دارد. برای استخراج جملات و کلمات از متن، از ابزار nltk بهره گرفته شده است. این ابزار شامل جداکننده های جملات و کلمات با استفاده از روش های متنوع می باشد. در این تحقیق، از روش Stanford Tokenizer که به عنوان یک قسمت از این ابزار در دسترس است، برای استخراج جملات و کلمات استفاده شده است. همچنین کلمات توقف و علائم نگارشی در متون شناسایی و حذف می شوند. به منظور ساخت گراف SWGraph، میزان تکرار کلمه در متن استخراج و گراف های مورد نیاز در قالب فایل ذخیره شده اند. برخی اطلاعات دیگر مانند کلیدواژه ها و گراف های نحوی استخراج می شوند.

۵-۵- ابرپارامترها

در سیستم‌های هوش مصنوعی، به ویژه در سیستم‌های یادگیری عمیق، برخی از پارامترها به صورت پیش‌فرض تنظیم شده‌اند. این پارامترها شامل عواملی هستند که قبل از شروع فرآیند یادگیری تعیین می‌شوند، از جمله تعداد لایه‌های شبکه، تعداد نوروها در هر لایه، ابعاد بردارهای جانشانی، تعداد داده‌های موجود در هر دسته کوچک^۱ برای یادگیری، تعداد تکرار^۲ ورود داده به شبکه، و غیره. این پارامترها به عنوان ابرپارامترها شناخته می‌شوند.

در این قسمت، از اطلاعات فراهم شده توسط مدل‌ها و تجربیات شخصی، جزئیات مربوط به ابرپارامترهای استفاده‌شده معرفی شده‌اند. در این پژوهش، تعداد دوره‌های ورود داده به شبکه برابر با ۲۰ در نظر گرفته شده است، اما برای جلوگیری از برازش اضافی مدل و به منظور بهینه‌سازی فرآیند یادگیری، از ماژول توقف زودهنگام^۳ استفاده گردیده است. همچنین، اندازه دسته به ۶۴ تنظیم شده است تا علاوه بر بهبود کارایی آموزشی، به بهینه‌سازی مصرف منابع محاسباتی نیز کمک کند. تعداد کلمات در هر جمله به ۵۰ کلمه محدود شده است، این موضوع مدل را قابل تطبیق با جملات بلند و پیچیده می‌سازد. اندازه خروجی هر ماژول به ۶۴ تنظیم شده است تا اطلاعات کلیدی را حفظ کرده و از ابعاد خروجی کاهش یابد. همچنین، تعداد حداکثر کلمات در هر جمله به ۱۰۰ محدود شده است تا دقت و کارایی مدل در تحلیل و خلاصه‌سازی متون بهبود یابد. نرخ یادگیری با مقدار ۰/۰۰۰۵ تنظیم شده است تا فرآیند آموزش به صورت پایدار و متوازن ادامه یابد. همچنین، تعداد ۴ جمله برگزیده جهت تولید خلاصه در نظر گرفته شده است تا اطلاعات کلیدی از متن مبدأ به درستی استخراج گردد. همچنین به منظور تنظیم میزان آستانه‌ی شباهت کسینوسی در ساختار SentSemanticGraph آزمایشات متنوعی انجام گرفت، که میزان ۰/۷ برای این ابرپارامتر در نظر گرفته شد. این تنظیمات با دقت و با استفاده از تجربیات به دست آمده، به بهبود و بهینه‌سازی یادگیری و تحلیل متون کمک نموده‌اند.

۵-۶- فرآیند یادگیری

در مسیر یادگیری مدل‌ها در این تحقیق، یک فرآیند یکپارچه از سه مرحله اصلی بهره‌مند شده است: آموزش، ارزیابی، و آزمون. در فرآیند آموزش، از مجموعه داده‌های آموزشی برای تنظیم وزن‌ها و پارامترهای مدل با استفاده از تابع خطای CrossEntropy و الگوریتم بهینه‌سازی Adam استفاده شده است. الگوریتم Adam به دلیل توانمندی آن در مسائل یادگیری عمیق انتخاب شده است. در

^۱ Batch

^۲ epoch

^۳ Early stopping

مرحله ارزیابی، از مجموعه داده ارزیابی برای سنجش عملکرد مدل استفاده شد. در این مرحله، معیارهای ارزیابی ROUGE-1، ROUGE-2، ROUGE-L و معیارهای دیگر مثل دقت و F1 محاسبه می‌شود. این ارزیابی به منظور انتخاب بهترین مدل برای مرحله آزمون اجرا می‌گردد. در مرحله آزمون، مدل با استفاده از داده‌های بخش آزمون که جدا از بخش‌های دیگر مجموعه داده است، ارزیابی می‌شود. این مرحله به‌عنوان یک ارزیابی نهایی و قابل اعتماد عمل می‌کند و نتایج نهایی و گزارشات مدل از این مرحله ارائه می‌شوند. همچنین درباره ابرپارامترها، از تنظیمات معینی که برخی از آن در بخش‌های قبلی معرفی شده‌اند، بهره گرفته شده است. به علاوه، در استفاده از مدل BERT، تنها لایه آخر برای یادگیری استفاده شده و لایه‌های دیگر از یادگیری خودداری شده‌اند.

۵-۷- نتایج و مقایسه با روش‌های دیگر

در این تحقیق، تنوع در رویکردها و ساختارهای مختلف در حوزه خلاصه‌سازی متن بر روی مجموعه داده‌ی CNN/DailyMail مورد بررسی و آزمایش قرار گرفته است. برخی از این آزمایش‌ها و رویکردها توانسته‌اند نتایج موفقیت‌آمیزی را در این زمینه به دنبال داشته باشند. جدول ۵ نتایج به‌دست‌آمده از معماری‌های متفاوت را بر روی مجموعه داده‌ی CNN/DailyMail نشان می‌دهد.

در این تحقیق، ابتدا ساختار SWGraph به‌عنوان ساختار پایه پیاده‌سازی شده و تلاش بر این بود که این ساختار بهبود یابد. همانطور که در جدول ۵ آمده است، با ترکیب لایه‌های KeywordGraph، SentSyntaxGraph، DocSemanticGraph و SentSemanticGraph با مدل SWGraph به صورت موازی، بهبودهای چشمگیری حاصل شده است. هر لایه ویژگی‌های خاص خود را داراست و می‌تواند اطلاعات مفیدی برای مسئله‌ی خلاصه‌سازی استخراج کند. در ترکیب‌های دوتایی، بهترین روش ترکیب موازی ساختار SentSemanticGraph با SWGraph مشاهده شده است. به همین دلیل، فرضیه‌ای دیگر آزمایش شده تا SentSemanticGraph و SWGraph به‌صورت توالی با هم قرار گیرند و آموزش ببینند. با وجود بهبود چشمگیر نسبت به ساختار SWGraph، این فرآیند نتوانست از ساختار مشابه ولی موازی خود بهترین نتیجه را برآورده کند.

به منظور بررسی ترکیب بخش‌های مختلف با هم، ترکیب SWGraph و SentSemanticGraph با بقیه‌ی ساختارهای گفته شده نیز آزمایش شد و نتایج آن در جدول ۵ قابل مشاهده است. مدل‌های سه‌تایی ذکر شده همگی نتایج موفق‌تری نسبت به مدل SWGraph داشته‌اند ولی نسبت به ترکیب SWGraph و SentSemanticGraph نتوانسته است بهبودی داشته باشد. با توجه به نتایج بدست آمده، می‌توان تحلیل‌هایی بدست آورد.

ساختار SentSemanticGraph به دلیل مشابه بودن معماری با DocSemanticGraph، همپوشانی زیادی دارند ولی SentSemanticGraph به ازای هر جمله متناسب با جمله‌ی خود بردار درست می‌کند، در حالی که DocSemanticGraph یک بردار برای تمامی جملات یک متن تولید می‌کند و از این لحاظ از SentSemanticGraph قدرت کم‌تری دارد. با ترکیب هر دوی آن‌ها با SWGraph، میزان پارامتر آموزشی و پیچیدگی یادگیری، افزایش پیدا می‌کند و نمی‌تواند بهبودی بیش‌تر حاصل کند.

ساختار SentSyntaxGraph، اطلاعات نحوی موجود در یک جمله را استخراج می‌کند. با توجه به نتایج بدست آمده، SentSemanticGraph که شامل لایه‌های یادگیری عمیق SageGraph و GraphAttentionPooling، می‌تواند اطلاعات نحوی و غنی‌تر از آن استخراج کند و به همین منظور در کنار هم نمی‌تواند بهبودی حاصل کند.

همچنین به دلیل استفاده از ساختارهای توجه در بخش‌های SWGraph و ساختار GraphAttentionPooling در داخل SentSemanticGraph، به کلمات کلیدی بیش‌تر توجه شده و اطلاعات خوبی بدست آورد. به همین منظور ساختار KeywordGraph نتوانست در کنار دو ساختار دیگر بهبودی حاصل کند. البته به دلیل استفاده از ابزار قدرتمند KeyBERT نتوانست نسبت به SWGraph بهبود خوبی بدست آورد.

ترکیب ساختار SWGraph با SentSemanticGraph بهترین مدل آزمایش شده در این پژوهش می‌باشد. به همین منظور ساختار SentSemanticGraph مدل نیز با جانشانی BERT¹، مدل کوچک‌تر^۱، آزمایش شد. که نتیجه‌ی آن نیز در جدول ۵ آمد. این نتایج نشان داد که جانشانی BERT قدرت خوبی در آموزش شبکه در مسئله‌ی خلاصه‌سازی خودکار متن، به مدل می‌دهد.

در فرآیند دستیابی به مدل پیشنهادی، آزمایشات زیادی با پارامترها و لایه‌های مختلف انجام گرفت که بهترین پیکره‌بندی آن در این پایان نامه گزارش شد. نتایج بدست آمده آزمایش‌های مختلف با اعمال تغییرات در بخش‌های مختلف مدل بدست آمده، در جدول ۶ نمایش داده شده است.

¹ Small BERT

جدول ۵: نتایج بدست آمده از ترکیب ساختارهای مختلف ارائه شده.

مدل	توضیحات	R1	R2	R3
SWGraph	جانشانی Glove، استفاده از CNN-LSTM و Position Embedding	42.35	19.34	38.82
SWGraph + KeywordGraph	جانشانی Glove، ۱۰ کلیدواژه استخراجی در قالب ساختار KeywordGraph به همراه SWGraph به صورت موازی	43.20	19.15	39.60
SWGraph + SentSyntaxGraph	جانشانی Glove، ساختار نحوی به همراه SWGraph به صورت موازی	43.32	19.10	39.52
SWGraph + DocSemanticGraph	جانشانی Glove، گراف معنایی در سطح سند به همراه SWGraph به صورت موازی	43.58	19.39	39.84
SWGraph + SentSemanticGraph (sequential)	جانشانی Glove، گراف معنایی در سطح جمله به همراه SWGraph به صورت توالی	43.44	19.25	39.59
SWGraph + SentSemanticGraph (parallel)	جانشانی Glove، گراف معنایی در سطح جمله به همراه SWGraph به صورت موازی	43.78	19.59	39.87
SWGraph + SentSemanticGraph (parallel and use BERT)	جانشانی BERT برای گراف معنایی در سطح جمله به همراه SWGraph به صورت موازی (جانشانی Glove برای SWGraph)	43.81	19.65	39.95
SWGraph + SentSemanticGraph + SentSyntaxGraph	جانشانی Glove، ترکیب ۳ ماژول SWGraph و SentSemanticGraph و SentSyntaxGraph به صورت موازی	43.59	19.41	39.66
SWGraph + SentSemanticGraph + KeywordGraph	جانشانی Glove، ترکیب ۳ ماژول SWGraph و SentSemanticGraph و KeywordGraph به صورت موازی	43.41	19.23	39.64
SWGraph + SentSemanticGraph + DocSemanticGraph	جانشانی Glove، ترکیب ۳ ماژول SWGraph و SentSemanticGraph و DocSemanticGraph به صورت موازی	43.59	19.45	39.78

جدول ۶: نتایج بدست آمده آزمایش‌های مختلف با اعمال تغییرات در بخش‌های مختلف مدل *SWGraph + SentSemanticGraph*.

حالات مختلف مدل <i>SWGraph</i> با <i>SentSemanticGraph</i> به صورت موازی و جانشانی کلمه <i>Glove</i> .	R1	R2	RL
استفاده از میانگین‌گیری به عنوان لایه‌ی ادغام به جای لایه‌ی <i>GraphAttentionPooling</i> در مدل <i>SentSemanticGraph</i>	43.23	19.17	39.32
استفاده از شبکه‌ی عصبی <i>GCN</i> به جای <i>SageGraph</i> در ساختار <i>SentSemanticGraph</i>	43.55	19.37	39.62
انتخاب ۳ جمله‌ی برتر به جای ۴ جمله در خلاصه‌ی خروجی	43.35	19.24	39.42
انتخاب ۵ جمله‌ی برتر به جای ۴ جمله در خلاصه‌ی خروجی	43.12	18.94	39.22
مدل نهایی	43.78	19.59	39.87

جدول ۷، نتایج بدست آمده از معروف‌ترین مدل‌ها در مسئله‌ی خلاصه‌سازی خودکار متن با رویکرد استخراجی به همراه بهترین مدل پیشنهادی در این پژوهش را نشان می‌دهد. همانطور که پیش‌تر ذکر شد *BERT* توانست موفقیت‌هایی را در زمینه‌ی خلاصه‌سازی بدست آورد. موارد ۱۰ تا ۱۸ جدول ۷، همگی از *BERT* استفاده نموده اند. *Small BERT*، *Large BERT* و *Roberta* از انواع *BERT* هستند.

همانطور که در جدول ۷ آمده است، مدل پیشنهادی با *Glove* توانسته است بهترین نتایج را در بین مدل‌های غیر *BERT* بدست آورد. همچنین مدل پیشنهادی با *BERT* نیز در بین مدل‌های همراه با *Small BERT* نیز توانسته بسیار موفق ظاهر شده و نتایج بسیار خوبی را بدست آورد. مدل پیشنهادی به دلیل تعداد پارامتر کمتر و عدم استفاده از *BERT*‌های قوی‌تر، نتوانسته است از مدل‌های مبتنی بر *Large BERT* و *Roberta* پیشی بگیرد.

مدل پیشنهادی در خلاصه‌سازی متون به بهبود چشمگیری نسبت به مدل *SWGraph* منجر شده است، نشان دهنده توانایی برتر این ترکیب در استخراج خلاصه مفید و کارآمد از متون می‌باشد.

مدل پیشنهادی با استفاده از ساختار *SWGraph* و ترکیب آن با *SentSemanticGraph* به نظر می‌رسد که نتایج قابل قبولی در خلاصه‌سازی متون حاصل کرده است. این ترکیب باعث بهبود چشمگیری نسبت به مدل *SWGraph* ایجاد شده و در مقایسه با ترکیب‌های دیگر نتایج خوبی داشته است.

جدول ۷: نتایج مدل‌های معروف در حوزه‌ی خلاصه‌سازی خودکار متن با رویکرد استخراجی و مقایسه آن‌ها با مدل پیشنهادی.

ردیف	مدل	R-1	R-2	R-L
۱	Oracle [72]	55.61	32.84	51.88
۲	SummaRuNNer [50]	39.60	16.20	35.30
۳	SUMO [78]	41.00	18.40	37.20
۴	LatentExtSum [66]	41.05	18.77	37.54
۵	BanditSum [63]	41.50	18.70	37.60
۶	NeuSum [56]	41.59	19.01	37.98
۷	JECS [79]	41.70	18.50	37.90
۸	HSG [51]	42.95	19.76	39.23
۹	DeepSum [74]	43.30	19.00	38.90
۱۰	BERT-Sum [47] (small)	43.25	20.24	39.63
۱۱	BERT-Sum [47] (large)	43.85	20.34	39.90
۱۲	Disco-BERT [68]	43.77	20.85	40.67
۱۳	TopicGraphSum [72] (large)	44.02	20.81	40.55
۱۴	MultiGras [70]	43.16	20.14	39.49
۱۵	HIBERT [67] (large)	43.19	20.46	39.72
۱۶	HIStruct [80] (Roberta)	43.65	20.54	40.03
۱۷	MatchSum [81] (Roberta)	44.41	20.86	40.55
۱۸	HAHSum [71] (large)	44.68	21.30	40.75
۱۹	SWGraph + SentSemanticGraph (Glove)	43.78	19.59	39.87
۲۰	SWGraph + SentSemanticGraph (small Bert)	43.81	19.65	39.95

۵-۸- جمع‌بندی

در این پژوهش، ابتدا به جزئیات پیاده‌سازی مدل پیشنهادی پرداخته شد. سپس معرفی و توضیح مجموعه داده CNN/DailyMail به عنوان مجموعه داده اصلی آزمایشات صورت گرفت. نحوه‌ی پیش‌پردازش داده‌ها، اعمال ابرپارامترها و فرآیند یادگیری در این مدل‌ها نیز به تفصیل مورد بررسی قرار گرفت. ابرپارامترهای استفاده شده در این پژوهش، شامل پارامترهایی مانند تعداد لایه‌ها، ابعاد بردارهای جانشانی، و نرخ یادگیری برای هر مدل بودند. پس از تعیین این ابرپارامترها، فرایند یادگیری مدل‌ها با استفاده از مجموعه داده مذکور اجرا شد.

در بخش نتایج، هر یک از ساختارها به طور جداگانه آزموده و با سایر روش‌ها مقایسه شد. ترکیب SWGraph با SentSemanticGraph به عنوان بهترین ساختار معرفی شد که بهبود چشمگیری نسبت به ساختارهای دیگر ارائه داد. همچنین، این ترکیب با جانشانی BERT نیز آموزش داده و آزموده شد. آزمایشات روی ساختارهای متنوع و مختلف نشان داد که معماری‌های گوناگون تأثیر زیادی بر کارایی سیستم دارند. همچنین، در ساختار SentSemanticGraph، استفاده از SageGraph به جای GCN و استفاده از GraphAttentionPooling به جای ادغام میانگیری به عنوان لایه‌های یادگیری عمیق، بهبودهای قابل توجهی در نتایج به دنبال داشت.

مدل پیشنهادی در مقایسه با مدل‌های دیگر توانست موفقیت‌های خوبی بدست آورد و در بین مدل‌های خلاصه‌سازی استخراجی متن، جایگاه مناسبی بدست آورد.

فصل ششم

نتیجه گیری و کارهای آینده

۶-۱- مقدمه

۶-۲- نتیجه گیری

در این تحقیق، تأثیر دیدگاه‌ها و ارتباطات مختلف در خلاصه‌سازی خودکار متن مورد بررسی قرار گرفت و انواع ساختارها مورد ارزیابی قرار گرفتند. نتایج آزمایش‌ها نشان دادند که استفاده از گراف‌ها و مدل‌سازی روابط و دیدگاه‌ها بهبودهای قابل توجهی در استخراج ویژگی‌های جملات متن دارد، که منجر به تولید خلاصه‌های کیفی با توجه به مفاهیم زمینه‌ای متن شده است. ارتباطات نحوی و معنایی نیز در خلاصه‌سازی خودکار متن تأثیرگذار بوده و روش‌های مبتنی بر کلیدواژه و عناوین نیز برای بهبود عملکرد استفاده شده‌اند. همچنین، گراف به عنوان یک ساختار داده مؤثر در مدل‌کردن روابط اثرگذار ثابت شده و جانشانی BERT به عنوان یک روش قدرتمند در استخراج خلاصه شناخته شد. مدل پیشنهادی نیز نتوانست جایگاه خوبی در میان مدل‌های خلاصه‌سازی متن به دست آورد و بهترین عملکرد را ارائه نمود.

در این پژوهش، به بررسی تأثیر دیدگاه‌ها و روابط مختلف در یک متن به جهت خلاصه‌سازی خودکار متن پرداخته شد. ساختارهای مختلف در این پژوهش ارائه و تأثیر هر کدام مورد بررسی قرار

گرفت. نتایج آزمایش‌های مختلف بر روی سیستم‌های توسعه داده شده، نشان داد که استفاده از گراف‌ها و مدل‌سازی انواع روابط و دیدگاه‌ها به وسیله‌ی آن می‌تواند ویژگی‌های مناسبی برای جملات یک متن استخراج کرده و خلاصه‌ی بهتری با توجه به مفاهیم زمینه‌ای آن متن تولید کند. روابط نحوی و معنایی در متن بر روی خلاصه‌سازی خودکار متن، توانست بهبودهای مناسبی ایجاد کند. همچنین از روش‌های مبتنی بر کلیدواژه و عناوین نیز به منظور بهبود خلاصه‌سازی استفاده شد و توانست ویژگی‌های مناسبی برای استخراج خلاصه بدست آورد. نتایج بدست آمده حاکی از آن است که در خلاصه‌سازی خودکار متن، روابط و دیدگاه‌های مختلف در یک متن می‌تواند مفید واقع شوند. همچنین گراف نیز به عنوان یک ساختار داده می‌تواند این روابط را به خوبی مدل کند. همچنین تغییر پارامترها، پیش‌پردازش‌ها و لایه‌ها می‌تواند بر روی استخراج خلاصه از متن تأثیرات مهمی بگذارند. در این پژوهش، جانشرانی BERT به عنوان یک جانشرانی قدرتمند در خلاصه‌سازی شناخته شد. همچنین مدل پیشنهادی توانست جایگاه خوبی در بین مدل‌های خلاصه‌سازی استخراجی متن بدست آورد و جزو بهترین‌ها قرار گیرد. نتایج نشان داد این روش به بهبود خلاصه‌سازی خودکار متن منجر می‌شود.

۳-۶- کارهای آینده

با توجه به اهمیت روابط بین جملات در فرآیند خلاصه‌سازی، پژوهش‌های آینده می‌توانند به بررسی دقیق‌تر و ساخت گراف با گره‌های جملات و ارتباطات بین آن‌ها با استفاده از شبکه‌های عصبی گراف تمرکز کنند.

استفاده از چندین تابع خطا و آموزش مسائل جانبی می‌تواند در فرآیند آموزش شبکه‌های عصبی برای مسئله‌ی خودکار خلاصه‌سازی متن اثربخش باشد. به عنوان مثال، اضافه کردن یک دسته‌بند کلمات برای استخراج کلمات حاضر در خلاصه، به همراه دیگر شبکه‌های عصبی، می‌تواند بهبود معنایی مدل را فراهم کند.

از طرفی، یکی از ضعف‌های مدل‌های خلاصه‌سازی استخراجی متن، ثابت بودن تعداد جملات استخراج شده است. اگر یک مدل بتواند تشخیص دهد چند جمله از متن ورودی باید استخراج شود و همچنین پراکندگی موضوع در جملات مناسب باشد، می‌توان خلاصه‌های مفیدتری تولید کرد. به همین منظور می‌توان از روش‌های خوشه‌بندی استفاده نمود.

روابط بسیاری می‌توان از یک متن استخراج کرد، چند مورد از این روابط در این پژوهش مورد بررسی قرار گرفت. می‌توان در ادامه به بررسی انواع روابط در بین موجودیت‌های یک متن و تأثیر هر کدام از آن‌ها، پرداخته شود.

منابع و مأخذ

- [1] T. Falke, "Automatic Structured Text Summarization with Concept Maps." Darmstadt, Apr. 12, 2019. [Online]. Available: <https://tuprints.ulb.tu-darmstadt.de/id/eprint/8430>
- [2] W. S. El-Kassas, C. R. Salama, A. A. Rafea, and H. K. Mohamed, "Automatic text summarization: A comprehensive survey," *Expert Syst. Appl.*, vol. 165, p. 113679, Mar. 2021, doi: 10.1016/j.eswa.2020.113679.
- [3] P. Verma and A. Verma, "A Review on Text Summarization Techniques," *J. Sci. Res.* 641, 2020.
- [4] F. Kiyani and O. Tas, "A survey automatic text summarization," *Pressacademia*, vol. 5, no. 1, pp. 205–213, Jun. 2017, doi: 10.17261/Pressacademia.2017.591.
- [5] A. Vaswani *et al.*, "Attention Is All You Need," 2017, doi: 10.48550/ARXIV.1706.03762.
- [6] V. Stuart, "Biomedical Knowledge Discovery in Networks Through Language/Graphical Models and Machine Learning." Aug. 2018. [Online]. Available: <https://persagen.com/resources/biokdd-review.html>
- [7] A. Cohan *et al.*, "A Discourse-Aware Attention Model for Abstractive Summarization of Long Documents," 2018, doi: 10.48550/ARXIV.1804.05685.
- [8] R. Nallapati, B. Zhou, C. N. dos santos, C. Gulcehre, and B. Xiang, "Abstractive Text Summarization Using Sequence-to-Sequence RNNs and Beyond." arXiv, Aug. 26, 2016. Accessed: Aug. 03, 2023. [Online]. Available: <http://arxiv.org/abs/1602.06023>
- [9] M. Yasunaga, R. Zhang, K. Meelu, A. Pareek, K. Srinivasan, and D. Radev, "Graph-based Neural Multi-Document Summarization," in *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, Vancouver, Canada: Association for Computational Linguistics, 2017, pp. 452–462. doi: 10.18653/v1/K17-1045.
- [10] N. I. Nikolov, M. Pfeiffer, and R. H. R. Hahnloser, "Data-driven Summarization of Scientific Articles." arXiv, Apr. 24, 2018. Accessed: Aug. 03, 2023. [Online]. Available: <http://arxiv.org/abs/1804.08875>
- [11] K. M. Hermann *et al.*, "Teaching Machines to Read and Comprehend," 2015, doi: 10.48550/ARXIV.1506.03340.
- [12] T. Hastie, J. Friedman, and R. Tibshirani, *The Elements of Statistical Learning*. in Springer Series in Statistics. New York, NY: Springer New York, 2001. doi: 10.1007/978-0-387-21606-5.

- [13] J. Heaton, "Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning," *Genet. Program. Evolvable Mach.*, vol. 19, no. 1, pp. 305–307, Jun. 2018, doi: 10.1007/s10710-017-9314-z.
- [14] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.
- [15] M. Wiering and M. Van Otterlo, Eds., *Reinforcement Learning: State-of-the-Art*, vol. 12. in *Adaptation, Learning, and Optimization*, vol. 12. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. doi: 10.1007/978-3-642-27645-3.
- [16] Md. J.-U. Rahman, R. I. Sultan, F. Mahmud, A. Shawon, and A. Khan, *Ensemble of Multiple Models For Robust Intelligent Heart Disease Prediction System*. 2019. doi: 10.1109/CEEICT.2018.8628152.
- [17] V. Markapuri, G. LaVessi, R. Stewart, and D. Wagner, *Bombus Species Image Classification*. 2020.
- [18] R. Jayawardana and T. Bandaranayake, *ANALYSIS OF OPTIMIZING NEURAL NETWORKS AND ARTIFICIAL INTELLIGENT MODELS FOR GUIDANCE, CONTROL, AND NAVIGATION SYSTEMS*. 2021.
- [19] R. Zaheer and H. Shaziya, "A Study of the Optimization Algorithms in Deep Learning," in *2019 Third International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, India: IEEE, Jan. 2019, pp. 536–539. doi: 10.1109/ICISC44355.2019.9036442.
- [20] K. Janocha and W. M. Czarnecki, "On Loss Functions for Deep Neural Networks in Classification," 2017, doi: 10.48550/ARXIV.1702.05659.
- [21] J.-C. Vialatte, "On convolution of graph signals and deep learning on graph domains," 2018.
- [22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
- [23] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [24] P. Kim, "Convolutional Neural Network," in *MATLAB Deep Learning*, Berkeley, CA: Apress, 2017, pp. 121–147. doi: 10.1007/978-1-4842-2845-6_6.
- [25] S. Gupta, "Deep learning based human activity recognition (HAR) using wearable sensor data," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 2, p. 100046, Nov. 2021, doi: 10.1016/j.jjime.2021.100046.
- [26] Chaitanya Bharathi Institute of Technology(Autonomous), K. M.Tarwani, and S. Edem, "Survey on Recurrent Neural Network in Natural Language Processing," *Int. J. Eng. Trends Technol.*, vol. 48, no. 6, pp. 301–304, Jun. 2017, doi: 10.14445/22315381/IJETT-V48P253.

- [27] Z. Cui, R. Ke, Z. Pu, and Y. Wang, "Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values," *Transp. Res. Part C Emerg. Technol.*, vol. 118, p. 102674, Sep. 2020, doi: 10.1016/j.trc.2020.102674.
- [28] V. Williams *et al.*, *Development of PPTNet a Neural Network for the Rapid Prototyping of Pulsed Plasma Thrusters*. 2019.
- [29] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997, doi: 10.1109/78.650093.
- [30] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to Sequence Learning with Neural Networks," in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2014. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2014/file/a14ac55a4f27472c5d894ec1c3c743d2-Paper.pdf
- [31] S. Tyagi, K. Chakraborty, V. Nunna, and A. Singh, "Finding Beauty Products Chemicals Using Token Identification and Entity Recognition Transformer Model," 2023. doi: 10.3233/AISE230014.
- [32] H. Wang, Y. Zhang, and X. Yu, "An Overview of Image Caption Generation Methods," *Comput. Intell. Neurosci.*, vol. 2020, pp. 1–13, Jan. 2020, doi: 10.1155/2020/3062706.
- [33] R. Patel and S. Patel, "Deep Learning for Natural Language Processing," Jan. 2019.
- [34] J. Cheng, L. Dong, and M. Lapata, "Long Short-Term Memory-Networks for Machine Reading," 2016, doi: 10.48550/ARXIV.1601.06733.
- [35] A. Bondy and U. S. R. Murty, "Graph theory: Graduate texts in mathematics," 2008.
- [36] A. L. Barabási, *Network Science*. Cambridge University Press, Cambridge. 2016.
- [37] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks." arXiv, Feb. 22, 2017. Accessed: Nov. 24, 2023. [Online]. Available: <http://arxiv.org/abs/1609.02907>
- [38] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive Representation Learning on Large Graphs," 2017, doi: 10.48550/ARXIV.1706.02216.
- [39] M. Zhang, G. Zhou, W. Yu, N. Huang, and W. Liu, "A Comprehensive Survey of Abstractive Text Summarization Based on Deep Learning," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–21, Aug. 2022, doi: 10.1155/2022/7132226.

- [40] H. Gao and S. Ji, "Graph U-Nets," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–1, 2021, doi: 10.1109/TPAMI.2021.3081010.
- [41] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed Representations of Words and Phrases and their Compositionality," 2013, doi: 10.48550/ARXIV.1310.4546.
- [42] J. Pennington, R. Socher, and C. Manning, "Glove: Global Vectors for Word Representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar: Association for Computational Linguistics, 2014, pp. 1532–1543. doi: 10.3115/v1/D14-1162.
- [43] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," 2018, doi: 10.48550/ARXIV.1810.04805.
- [44] M. Indu and K. V. Kavitha, "Review on text summarization evaluation methods," in *2016 International Conference on Research Advances in Integrated Navigation Systems (RAINS)*, Bangalore, India: IEEE, May 2016, pp. 1–4. doi: 10.1109/RAINS.2016.7764406.
- [45] M. Barbella and G. Tortora, "Rouge Metric Evaluation for Text Summarization Techniques," *SSRN Electron. J.*, 2022, doi: 10.2139/ssrn.4120317.
- [46] B. Mutlu, E. A. Sezer, and M. A. Akcayol, "Candidate sentence selection for extractive text summarization," *Inf. Process. Manag.*, vol. 57, no. 6, p. 102359, Nov. 2020, doi: 10.1016/j.ipm.2020.102359.
- [47] Y. Liu and M. Lapata, "Text Summarization with Pretrained Encoders," 2019, doi: 10.48550/ARXIV.1908.08345.
- [48] Y. Xu and M. Lapata, "Text Summarization with Oracle Expectation," 2022, doi: 10.48550/ARXIV.2209.12714.
- [49] M. Mohd, R. Jan, and M. Shah, "Text document summarization using word embedding," *Expert Syst. Appl.*, vol. 143, p. 112958, Apr. 2020, doi: 10.1016/j.eswa.2019.112958.
- [50] R. Nallapati, F. Zhai, and B. Zhou, "SummaRuNNer: A Recurrent Neural Network based Sequence Model for Extractive Summarization of Documents," 2016, doi: 10.48550/ARXIV.1611.04230.
- [51] D. Wang, P. Liu, Y. Zheng, X. Qiu, and X. Huang, "Heterogeneous Graph Neural Networks for Extractive Document Summarization," 2020, doi: 10.48550/ARXIV.2004.12393.
- [52] A. Bichi, P. Keikhosrokiani, R. Hassan, and K. Almekhlafi, "Graph-Based Extractive Text Summarization Models: A Systematic Review," *J. Inf. Technol. Manag.*, vol. 14, no. 5th International Conference of Reliable Information and

Communication Technology (IRICT 2020), Jan. 2022, doi: 10.22059/jitm.2022.84899.

[53] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A Comprehensive Survey on Graph Neural Networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021, doi: 10.1109/TNNLS.2020.2978386.

[54] J. Zhou *et al.*, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020, doi: 10.1016/j.aiopen.2021.01.001.

[55] H. Saggion and T. Poibeau, "Automatic Text Summarization: Past, Present and Future," in *Multi-source, Multilingual Information Extraction and Summarization*, T. Poibeau, H. Saggion, J. Piskorski, and R. Yangarber, Eds., in *Theory and Applications of Natural Language Processing.*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 3–21. doi: 10.1007/978-3-642-28569-1_1.

[56] Q. Zhou, N. Yang, F. Wei, S. Huang, M. Zhou, and T. Zhao, "Neural Document Summarization by Jointly Learning to Score and Select Sentences," 2018, doi: 10.48550/ARXIV.1807.02305.

[57] Rahul, S. Adhikari, and Monika, "NLP based Machine Learning Approaches for Text Summarization," in *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India: IEEE, Mar. 2020, pp. 535–538. doi: 10.1109/ICCMC48092.2020.ICCMC-00099.

[58] K. Kandasamy and P. Koroth, "An integrated approach to spam classification on Twitter using URL analysis, natural language processing and machine learning techniques," in *2014 IEEE Students' Conference on Electrical, Electronics and Computer Science*, Bhopal: IEEE, Mar. 2014, pp. 1–5. doi: 10.1109/SCEECS.2014.6804508.

[59] G. Silva *et al.*, "Automatic Text Document Summarization Based on Machine Learning," in *Proceedings of the 2015 ACM Symposium on Document Engineering*, Lausanne Switzerland: ACM, Sep. 2015, pp. 191–194. doi: 10.1145/2682571.2797099.

[60] T. Jo, "K nearest neighbor for text summarization using feature similarity," in *2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*, Khartoum, Sudan: IEEE, Jan. 2017, pp. 1–5. doi: 10.1109/ICCCCEE.2017.7866705.

[61] M. A. Fattah, "A hybrid machine learning model for multi-document summarization," *Appl. Intell.*, vol. 40, no. 4, pp. 592–600, Jun. 2014, doi: 10.1007/s10489-013-0490-0.

[62] D. Miller, "Leveraging BERT for Extractive Text Summarization on Lectures," 2019, doi: 10.48550/ARXIV.1906.04165.

- [63] Y. Dong, Y. Shen, E. Crawford, H. van Hoof, and J. C. K. Cheung, "BanditSum: Extractive Summarization as a Contextual Bandit," 2018, doi: 10.48550/ARXIV.1809.09672.
- [64] A. See, P. J. Liu, and C. D. Manning, "Get To The Point: Summarization with Pointer-Generator Networks," 2017, doi: 10.48550/ARXIV.1704.04368.
- [65] A. M. Rush, S. Chopra, and J. Weston, "A Neural Attention Model for Abstractive Sentence Summarization." arXiv, Sep. 03, 2015. Accessed: Dec. 13, 2023. [Online]. Available: <http://arxiv.org/abs/1509.00685>
- [66] X. Zhang, M. Lapata, F. Wei, and M. Zhou, "Neural Latent Extractive Document Summarization," 2018, doi: 10.48550/ARXIV.1808.07187.
- [67] X. Zhang, F. Wei, and M. Zhou, "HIBERT: Document Level Pre-training of Hierarchical Bidirectional Transformers for Document Summarization," 2019, doi: 10.48550/ARXIV.1905.06566.
- [68] J. Xu, Z. Gan, Y. Cheng, and J. Liu, "Discourse-Aware Neural Extractive Text Summarization," 2019, doi: 10.48550/ARXIV.1910.14142.
- [69] C. Mallick, A. K. Das, M. Dutta, A. K. Das, and A. Sarkar, "Graph-Based Text Summarization Using Modified TextRank," in *Soft Computing in Data Analytics*, vol. 758, J. Nayak, A. Abraham, B. M. Krishna, G. T. Chandra Sekhar, and A. K. Das, Eds., in *Advances in Intelligent Systems and Computing*, vol. 758. , Singapore: Springer Singapore, 2019, pp. 137–146. doi: 10.1007/978-981-13-0514-6_14.
- [70] B. Jing, Z. You, T. Yang, W. Fan, and H. Tong, "Multiplex Graph Neural Network for Extractive Text Summarization," 2021, doi: 10.48550/ARXIV.2108.12870.
- [71] R. Jia, Y. Cao, H. Tang, F. Fang, C. Cao, and S. Wang, "Neural Extractive Summarization with Hierarchical Attentive Heterogeneous Graph Network," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online: Association for Computational Linguistics, 2020, pp. 3622–3631. doi: 10.18653/v1/2020.emnlp-main.295.
- [72] P. Cui, L. Hu, and Y. Liu, "Enhancing Extractive Text Summarization with Topic-Aware Graph Neural Networks," 2020, doi: 10.48550/ARXIV.2010.06253.
- [73] R. C. Belwal, S. Rai, and A. Gupta, "A new graph-based extractive text summarization using keywords or topic modeling," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 10, pp. 8975–8990, Oct. 2021, doi: 10.1007/s12652-020-02591-x.
- [74] A. Joshi, E. Fidalgo, E. Alegre, and L. Fernández-Robles, "DeepSumm: Exploiting topic models and sequence to sequence networks for extractive text summarization," *Expert Syst. Appl.*, vol. 211, p. 118442, Jan. 2023, doi: 10.1016/j.eswa.2022.118442.

- [75] R. Yuan, S. Sun, Z. Wang, Z. Cao, and W. Li, “Separating Context and Pattern: Learning Disentangled Sentence Representations for Low-Resource Extractive Summarization,” in *Findings of the Association for Computational Linguistics: ACL 2023*, Toronto, Canada: Association for Computational Linguistics, 2023, pp. 7575–7586. doi: 10.18653/v1/2023.findings-acl.479.
- [76] S. Takase and N. Okazaki, “Positional Encoding to Control Output Sequence Length,” in *Proceedings of the 2019 Conference of the North*, Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 3999–4004. doi: 10.18653/v1/N19-1401.
- [77] Q. Zhong, L. Ding, J. Liu, B. Du, H. Jin, and D. Tao, “Knowledge Graph Augmented Network Towards Multiview Representation Learning for Aspect-Based Sentiment Analysis,” *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 10, pp. 10098–10111, Oct. 2023, doi: 10.1109/TKDE.2023.3250499.
- [78] Y. Liu, I. Titov, and M. Lapata, “Single Document Summarization as Tree Induction,” in *Proceedings of the 2019 Conference of the North*, Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 1745–1755. doi: 10.18653/v1/N19-1173.
- [79] J. Xu and G. Durrett, “Neural Extractive Text Summarization with Syntactic Compression,” 2019, doi: 10.48550/ARXIV.1902.00863.
- [80] Q. Ruan, M. Ostendorff, and G. Rehm, “HiStruct+: Improving Extractive Text Summarization with Hierarchical Structure Information,” 2022, doi: 10.48550/ARXIV.2203.09629.
- [81] M. Zhong, P. Liu, Y. Chen, D. Wang, X. Qiu, and X. Huang, “Extractive Summarization as Text Matching,” 2020, doi: 10.48550/ARXIV.2004.08795.

Abstract:

Automatic text summarization is an active area in natural language processing research that has brought significant improvements in the field by utilizing deep neural networks and graphs. In this study, an extractive approach to text summarization is considered, employing semantic and syntactic graphs to enhance the accuracy and quality of summaries. The research also aims to investigate the impact of different perspectives and relationships within the text on extractive summarization.

The problem of summarization, as a crucial task in natural language processing, seeks to generate meaningful summaries of lengthy texts and documents. In this regard, deep neural networks are employed as the primary tool for extracting important features from texts. By using semantic and syntactic graphs, relationships between words and sentences, as well as semantic information and better structures, can be extracted. These graphs are constructed based on semantic relationships between words, grammatical structures of sentences, and various relationships between the entities present in the text. Additionally, a heterogeneous graph is used in this research for document modeling, attempting to extract various types of information using different methods and perspectives in text summarization.

The model learns through supervised training using training data. By employing a loss function based on the expected output, the model is improved, and network weights are updated. The conducted evaluations indicate a significant increase in conveying meaning and preserving key information in the summaries. The combination of deep neural networks with graphs and the utilization of different perspectives and relationships in the summarization domain, as an innovative approach, demonstrate successful results and lead to a considerable enhancement in the quality and accuracy of summaries. This research highlights that adopting this approach results in clear progress in the field of automatic text summarization.

Keywords: 1. Extractive Text Summarization 2. Deep Neural Networks 3. Semantic Graphs 4. Syntactic Graphs 5. Natural Language Processing 6. Graph Neural Networks 7. Title-based Summarization



Vice Chancellor for Research and Technology

Department of Artificial intelligence

Faculty of Computer engineering

University of Isfahan

A thesis entitled

Extractive text summarization using deep learning approaches

was submitted by

AmirReza Seddighin



University of Isfahan

Faculty of Computer engineering

Department of Artificial intelligence and robotics

Master of Science Thesis

Extractive text summarization using deep learning approaches

By:

AmirReza Seddighin

Supervisor:

Dr. HamidReza Baradaran Kashani

Advisor:

Dr. Afsaneh Fatemi