



دانشگاه صنعتی امیرکبیر

(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر

پروژه تحقیقاتی درس سمینار

ماشین تورینگ عصبی

نگارش

زهرا زنجانی

استاد درس

دکتر رضا صفابخش

۱۴۰۱ - ۱۴۰۲

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

چکیده

خلاصه سازی نقش مهمی در علم اطلاعات و بازیابی دارد، زیرا ارتباط نزدیکی با فشرده سازی داده ها و درک اطلاعات دارد. توانایی تولید خلاصه های مناسب می تواند موجب بهبود کارآمدی سیستم های استخراج اطلاعات و صرفه جویی در وقت انسان ها شود. خلاصه سازی خودکار به عنوان یک کار برجسته در پردازش زبان طبیعی¹ ظاهر شده است. با این حال، علیرغم اهمیت آن، زمینه خلاصه سازی خودکار تا حد زیادی حل نشده باقی مانده است. این گزارش مروری جامع از وضعیت فعلی خلاصه سازی خودکار ارائه می کند و رویکردها، تکنیک ها و معیارهای ارزیابی مختلف به کار گرفته شده در این زمینه را بررسی می کند.

واژه های کلیدی:

خلاصه سازی متن، پردازش زبان طبیعی، یادگیری عمیق

¹natural language processing (NLP)

فهرست مطالب

صفحه

عنوان

۱	مقدمه	۱
۳	روش‌های مبتنی بر ساختار	۲
۴	۱-۲ روش مبتنی بر درخت	
۴	۲-۲ روش مبتنی بر قالب	
۵	۳-۲ روش مبتنی بر هستان شناسی	
۵	۴-۲ روش عبارت مقدمه و بدنه	
۶	۵-۲ روش مبتنی بر گراف	
۶	۶-۲ روش مبتنی بر قانون	
۷	روش‌های مبتنی بر شبکه‌ی عصبی	۳
۸	۱-۳ روش‌های مبتنی بر مدل کدگذار-کدگشا	
۹	۲-۳ روش‌های مبتنی بر مدل ترنسفورمرها	
۱۰	۳-۳ روش‌های مبتنی بر مدل‌های از پیش آموزش دیده	
۱۰	۴-۳ ایده‌های ارائه شده بهبود خلاصه سازی متون طولانی	
۱۲	روش‌های مبتنی بر یادگیری تقویتی	۴
۱۳	۱-۴ یادگیری تقویتی برای حل مسائل عمیق توالی به دنباله	
۱۳	۲-۴ یادگیری تقویتی برای ترکیب خلاصه‌های استخراجی و انتزاعی	
۱۴	۳-۴ یادگیری تقویتی برای ایجاد معیارها و پاداش‌های جدید	
۱۵	نتایج	۵
۱۶	جمع‌بندی	۶
۱۷	منابع و مراجع	
۲۰	واژه‌نامه‌ی فارسی به انگلیسی	
۲۳	واژه‌نامه‌ی انگلیسی به فارسی	

شکل	فهرست اشکال	صفحه
۱-۳	معماری پایه‌ی مدل کدگذار-کدگشا [؟]	۸
۲-۳	معماری پایه‌ی مدل سلسله مراتبی متغیر برای خلاصه‌سازی متقابل زبانی [۱۰]	۹

صفحه

فهرست جداول

جدول

فهرست اختصارات

عنوان اختصاری	عنوان کامل
ماتع	ماشین تورینگ عصبی
ماتع تکاملی	ماشین تورینگ عصبی تکاملی
ماتع ابر تکاملی	ماشین تورینگ عصبی ابر تکاملی
ماتع متوجه	ماشین تورینگ عصبی متوجه
ماتع پویا	ماشین تورینگ عصبی پویا
تکتوت	تکامل عصبی توپولوژی‌های تقویت‌کننده
ابرتوت	ابر تکامل عصبی توپولوژی‌های تقویت‌کننده
شبکه تات	شبکه تولید الگوی ترکیبی

فصل اول

مقدمه

با رشد روزافزون اینترنت، محتوای متنی در اینترنت (به عنوان مثال وب سایت‌ها، نظرات کاربران، اخبار، وبلاگ‌ها، شبکه‌های رسانه‌های اجتماعی و غیره) به صورت تصاعدی افزایش می‌یابد. در نتیجه، کاربران زمان زیادی را صرف یافتن اطلاعات مورد نظر خود می‌کنند و حتی نمی‌توانند تمام محتوای متنی نتایج جستجو را بخوانند و درک کنند. خلاصه‌سازی خودکار اسناد می‌تواند به شناسایی مهم‌ترین اطلاعات، ارائه خلاصه‌ای جامع و صرفه‌جویی در وقت خوانندگان کمک کند. خلاصه‌سازی خودکار متن فرآیند تولید یک متن کوتاه است که بخش‌های اصلی یک سند طولانی‌تر را پوشش می‌دهد. یک خلاصه خوب جنبه‌های مهمی مانند خوانایی، انسجام، نحو، غیر زائد بودن، ترتیب جملات، مختصر بودن، تنوع اطلاعات و پوشش اطلاعات را در نظر می‌گیرد [۳].

در سال‌های گذشته تلاش‌های زیادی برای تولید خلاصه‌سازی خودکار قابل قبول و خوانا صورت گرفته است. پژوهش‌های مرتبط با عمل خلاصه‌سازی خودکار متن در دهه ۵۰ میلادی شکل گرفتند. در یکی از این پژوهش‌ها لوهن و همکاران روشی برای خلاصه‌سازی اسناد علمی ارائه دادند که در آن تابعی بر اساس فرکانس تکرار کلمات یا عبارات به عنوان ویژگی تعریف می‌شود و با یادگیری وزن‌های مرتبط با این ویژگی‌ها خلاصه استخراج می‌شود [۱۱]. در کارهای تحقیقاتی اولیه، مدل‌های غیرعصبی مبتنی بر ساختار برای تولید خلاصه‌سازی خودکار مورد استفاده قرار گرفتند. با شروع دوره‌ی شبکه‌های عصبی عمیق پژوهش‌ها بر روی خلاصه‌سازی بیشتر شد. رویکردهای نوین خلاصه‌سازی شامل شبکه‌های عصبی عمیق دنباله به دنباله^۱، روش‌های بر پایه‌ی مدل تبدیل‌کننده^۲ و مدل‌های زبانی از پیش آموزش دیده^۳ می‌باشد. همچنین برخی از پژوهش‌های اخیر نشان داده‌اند، استفاده از رویکردهای مبتنی بر یادگیری تقویتی^۴ می‌تواند موجب بهبود معیارهای مختلف، از جمله امتیازات روز، کیفیت کلی، خوانایی، انسجام، نحو، غیر افزونگی، ترتیب جملات، مختصر بودن، تنوع اطلاعات، پوشش اطلاعات شود.

¹ Deep neural sequence to sequence models

² transformer

³ Pretrained language models (PTLMs)

⁴ reinforcement learning (RL)

فصل دوم

روش‌های مبتنی بر ساختار

روش‌های خلاصه‌سازی مبتنی بر ساختار طیفی از رویکردها را در بر می‌گیرند که از ویژگی‌های ساختاری متن ورودی برای تولید خلاصه‌های مختصر و منسجم استفاده می‌کنند. در این رویکرد اطلاعات مهم متن به یک ساختار از پیش تعریف شده داده می‌شود و خلاصه با توجه به ساختار ایجاد می‌شود. در این فصل روش‌های مبتنی بر درخت^۱، مبتنی بر قالب^۲، مبتنی بر هستان‌شناسی^۳، عبارت مقدمه و بدنه^۴، مبتنی بر گراف^۵ و مبتنی بر قانون^۶ مورد بررسی قرار می‌گیرد.

۱-۲ روش مبتنی بر درخت

روش مبتنی بر درخت در خلاصه‌سازی متن شامل استفاده از درخت‌های وابستگی برای نمایش سند متنی است. متن مبدأ ابتدا به درخت‌های وابستگی تبدیل می‌شود، که سپس در یک درخت واحد ادغام می‌شوند. سپس این درخت وابستگی ادغام شده به جمله ای تبدیل می‌شود که به عنوان جمله ترکیبی شناخته می‌شود. فرآیند تبدیل درخت وابستگی به رشته ای از کلمات را خطی سازی درخت می‌گویند. عملکرد این روش به انتخاب تجزیه کننده و حفظ وابستگی بین کلمات بستگی دارد. تکنیک‌های مختلفی پیشنهاد شده‌اند، مانند استفاده از تجزیه‌کننده‌های کم عمق برای ترکیب جملات مشابه، حذف زیردرخت‌های درخت‌های وابستگی برای فشرده‌سازی، و تولید درخت‌های تودرتو با استفاده از ساختارهای بلاغی و تجزیه وابستگی. به طور کلی، روش مبتنی بر درخت با استفاده از ساختار سند متنی، با هدف ایجاد خلاصه‌های مختصر و منسجم است. این روش به عملکرد تجزیه کننده‌ها وابسته است و این باعث محدود شدن کارایی می‌شود[۲].

۲-۲ روش مبتنی بر قالب

روش‌های مبتنی بر الگو در خلاصه‌سازی متن شامل استفاده از قالب‌های از پیش تعریف شده برای نمایش سند است. این قالب‌ها برای مطابقت با الگوها و قوانین خاص در محتوای متنی طراحی شده‌اند و امکان استخراج اطلاعات مرتبط را فراهم می‌کنند که می‌توان آن‌ها را در فضای قالب ترسیم کرد. این فرآیند شامل تطبیق متن با این الگوها و قوانین برای شناسایی محتوای متناسب با الگو است، که سپس محتوای خلاصه را نشان می‌دهد. این روش بسیار منسجم است زیرا خلاصه‌هایی تولید می‌کند که به ساختار و قالب قالب‌ها پایبند هستند. با این حال، یکی از چالش‌های پیش روی روش‌های مبتنی بر الگو، نیاز به تجزیه و تحلیل معنایی دقیق است، چرا که قالب‌ها نیاز به محتوای خاص و مرتبط برای پر

¹tree-based

²template-based

³ontology-based

⁴lead-and-body phrase

⁵graph-based

⁶rule-based

شدن دارند [۲].

۳-۲ روش مبتنی بر هستان شناسی

روش مبتنی بر هستی‌شناسی در خلاصه سازی متن شامل استفاده از پایگاه دانش یا هستی‌شناسی برای بهبود فرآیند خلاصه سازی است. این روش از این واقعیت بهره می‌برد که بسیاری از اسناد موجود در اینترنت به حوزه‌های خاصی با واژگان محدود مرتبط هستند که می‌توانند توسط هستی‌شناسی بهتر نمایش داده شوند. هستی‌شناسی نامگذاری و تعریف رسمی انواع موجودیت مربوط به یک دامنه خاص را ارائه می‌دهد که به عنوان پایگاه دانش عمل می‌کند. با استفاده از هستی‌شناسی، سیستم خلاصه سازی می‌تواند نمایش معنایی محتوای اطلاعات را بهبود بخشد و بسط پرس و جو را انجام دهد. تکنیک‌های مختلفی پیشنهاد شده است، مانند استفاده از هستی‌شناسی برای ساخت یک مدل معنایی، نگاشت جملات به گره‌های هستی‌شناسی، و محاسبه امتیاز مربوط به موجودیت برای رتبه بندی جملات. به طور کلی، روش مبتنی بر هستی‌شناسی از دانش خاص دامنه برای ایجاد خلاصه‌های دقیق‌تر و آموزنده‌تر استفاده می‌کند [۲]. لی و همکاران یک سیستم فازی را ارائه کرد که از هستی‌شناسی طراحی شده توسط متخصص حوزه اخبار استفاده می‌کند. جملات بر اساس طبقه بندی کننده اصطلاحی که از هستی‌شناسی استفاده می‌کند، طبقه بندی می‌شوند. مکانیزم استنتاج فازی درجه عضویت برای هر جمله را با توجه به طبقه بندی کننده اصطلاح براساس هستی‌شناسی دامنه محاسبه می‌کند [۹].

۴-۲ روش عبارت مقدمه و بدنه

روش عبارت مقدمه و بدنه یک رویکرد خلاصه‌سازی متن است که بر شناسایی و بازنگری جملات اصلی، معروف به جملات اصلی، در یک سند تمرکز دارد. این جملات اصلی معمولاً آموزنده هستند و خلاصه خوبی از محتوا ارائه می‌دهند. این روش شامل درج و جایگزینی عبارات در جمله اصلی برای ایجاد تجدید نظرهای معنایی مناسب است. با بازنویسی تکراری جمله اصلی، جملات خلاصه جدیدی تولید می‌شوند. با این حال، یکی از محدودیت‌های این روش این است که تجزیه می‌تواند عملکرد آن را کاهش دهد، و هیچ مدل تعمیم یافته‌ای برای خلاصه سازی وجود ندارد [۲]. ایشیکاوا و همکاران روش خلاصه سازی ترکیبی مبتنی بر روش فرکانس عبارت^۷ و عبارت مقدمه و بدنه پیشنهاد کردند. تابع توزیع زاویه‌ای ضربدر بسامد عبارت که وزن را به هر جمله برای شناسایی اهمیت اختصاص می‌دهد. دستورهای براساس اهمیت برای نوشتن خلاصه رتبه بندی می‌شوند [۷].

⁷Term frequency (TF)

۵-۲ روش مبتنی بر گراف

روش مبتنی بر نمودار یک رویکرد خلاصه سازی است که هر جمله در یک سند را به عنوان یک راس در یک نمودار نشان می دهد. جملات بر اساس روابط معنایی با یال ها به هم متصل می شوند و وزن یال ها نشان دهنده قدرت رابطه است. سپس از یک الگوریتم رتبه بندی نمودار برای تعیین اهمیت هر جمله در نمودار استفاده می شود. جملات با اهمیت بالاتر، که با وزن های بالاتر یا ارتباط بیشتر نشان داده می شوند، مهم تر در نظر گرفته می شوند و در خلاصه گنجانده می شوند. این روش نیازی به دانش عمیق زبانی یا حوزه ای ندارد و می تواند با انتخاب جملاتی با اهمیت بالا، خلاصه های مختصر و منسجمی ایجاد کند [۲]. مالپروس و اسکینیس از مرکزیت گره برای نشان دادن اهمیت یک اصطلاح در سند استفاده می کنند. مرکزیت های گره محلی و جهانی برای وزن دهی عبارت در نظر گرفته می شوند تا خلاصه را شکل دهند [۱۲].

۶-۲ روش مبتنی بر قانون

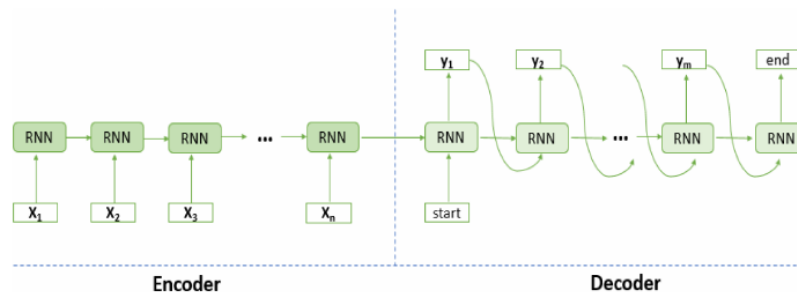
در این تکنیک، اسنادی که باید خلاصه شوند از نظر طبقات و فهرست جنبه ها به تصویر کشیده می شوند. ماژول انتخاب محتوا، مؤثرترین نامزد را از میان مواردی که توسط قوانین استخراج داده ایجاد می شود، انتخاب می کند تا به یک یا بسیاری از جنبه های یک دسته پاسخ دهد. در نهایت، الگوهای تولید برای تولید جملات طرح کلی استفاده می شود [۱۳].

فصل سوم

روش‌های مبتنی بر شبکه‌ی عصبی

۱-۳ روش‌های مبتنی بر مدل کدگذار-کدگشا

قبل از ظهور ترنسفورمرها، مدل‌های شبکه عصبی عمیق دنباله به دنباله بهترین مدل‌ها برای وظایف تولید متن از جمله ترجمه‌ی ماشینی و خلاصه‌سازی متن بوده‌اند. این مدل‌ها ورودی را از یک فرم به فرم دیگر نگاشت می‌کنند تا نتایج مورد نظر را تولید کنند. معماری کدگذار-کدگشا رویکرد اصلی برای مدل‌سازی مدل‌های دنباله به دنباله است. شکل ۱-۳ معماری پایه‌ی مدل کدگذار-کدگشا را شرح می‌دهد. شبکه‌های بازگشتی^۱ [۴] و حافظه‌های کوتاه مدت طولانی [۶] برای توالی طراحی شده‌اند و مناسب‌ترین معماری‌های یادگیری عمیق برای کدگذاری و پردازش داده‌های دنباله‌ای مانند متن هستند. اما این شبکه‌ها در مدیریت حافظه‌ی بلند مدت طولانی^۲ مشکل دارند.



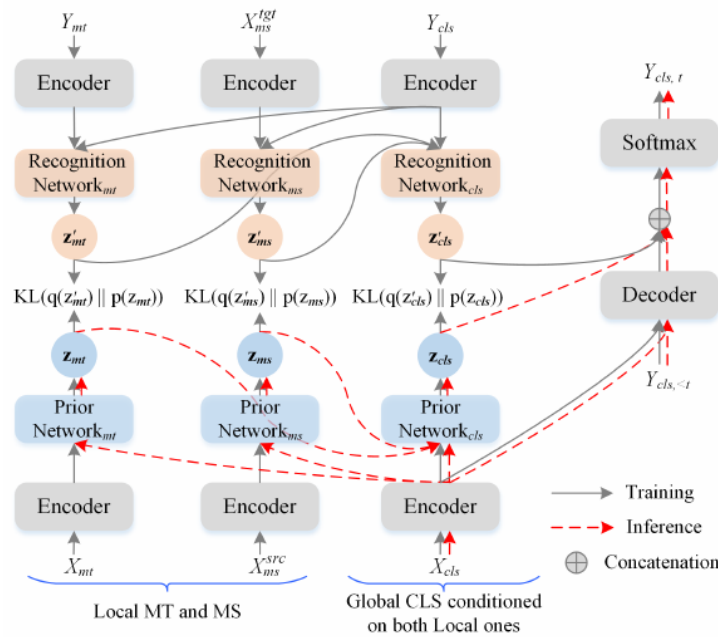
شکل ۱-۳: معماری پایه‌ی مدل کدگذار-کدگشا [۹]

یکی از مدل‌های کدگذار-کدگشای ارایه شده مدل سلسله مراتبی متغیر برای خلاصه‌سازی متقابل زبانی^۳ می‌باشد. مدل پیشنهادی شامل دو متغیر نهفته محلی، یکی برای ترجمه و دیگری برای خلاصه سازی، و یک متغیر نهفته جهانی برای خلاصه سازی بین زبانی است. متغیرهای پنهان محلی به ترتیب برای بازسازی ترجمه و خلاصه زبان مبدأ محدود می‌شوند. سپس از متغیر پنهان سراسری برای تولید خلاصه بین زبانی استفاده می‌شود. قسمت کدگذار دو بخش دارد که هر بخش وظیفه‌ی تولید یکی از متغیرهای پنهان محلی را دارد. بخش کدگشا با استفاده از نمایش‌های نهفته‌ی محلی خلاصه‌ی نهایی را تولید می‌کند. ساختار سلسله مراتبی مدل به آن اجازه می‌دهد تا رابطه سلسله مراتبی بین ترجمه، خلاصه سازی و خلاصه سازی بین زبانی را بیاموزد.

^۱recurrent neural network (RNN)

^۲long short-term memory networks(LSTM)

^۳cross-lingual



شکل ۲-۳: معماری پایه‌ی مدل سلسله مراتبی متغیر برای خلاصه‌سازی متقابل زبانی [۱۰]

متغیرهای محلی z_{mt} و z_{ms} به ترتیب برای ترجمه و خلاصه‌سازی طراحی شده‌اند. سپس z_{cls} جهانی برای خلاصه‌سازی بین زبانی است، خطوط خاکستری نشان‌دهنده فرآیند آموزشی است که مسئول تولید $(z'_{ms}, z'_{mt}, z'_{cls})$ از توزیع پسین متناظر پیش‌بینی شده توسط شبکه‌های شناسایی است که یادگیری شبکه‌های قبلی را هدایت می‌کند. خطوط قرمز چین نشان‌دهنده فرآیند استنتاج برای تولید نمایش‌های نهفته $(z_{cls}, z_{ms}, z_{mt})$ از توزیع‌های قبلی مربوطه پیش‌بینی شده توسط شبکه‌های قبلی است.

۲-۳ روش‌های مبتنی بر مدل ترنسفورمرها

با ظهور ترنسفورمرها [۱۶]، بهبودهای قابل توجهی در کیفیت نتایج خلاصه‌سازی خودکار به وجود آمد. ترنسفورمرها با استفاده از مکانیزم توجه به خود^۴ شباهت بین ورودی‌ها را بدون توجه به موقعیت موازی آن‌ها با حضور مستقل هر توکن در توالی ورودی مدل می‌کنند و به طور مؤثر مشکلات شبکه‌های بازگشتی را حل می‌کنند. یکی از جهت‌گیری‌های رایج پژوهشی، اصلاح یا تطبیق ترنسفورمرها و مدل‌های زبانی از پیش آموزش دیده با وظایف مختلف مانند خلاصه‌سازی است. مدل‌های مبتنی بر مدل‌های زبانی از پیش آموزش دیده که با هدف خلاصه‌سازی انتزاعی طراحی شده‌اند از ویژگی‌های معنایی و متنی غنی بازنمایی‌های زبان برای بهبود کیفیت و دقت خلاصه‌ها استفاده می‌کنند.

به عنوان مثال مدل پگاسوس^۵ [۱۷] یک مدل دنباله به دنباله کدگذار کدگشا مبتنی بر ترنسفورمر است که بر روی مجموعه‌های متنی بدون نظارت با هدف تولید جملات شکاف^۶ از قبل آموزش داده شده است. این مدل عملکرد مناسبی در خلاصه‌سازی متون کوتاه دارد. در حال حاضر بهترین مدل خلاصه‌سازی

^۴Self-attention

^۵PEGASUS

^۶gap sentences generation

متون کوتاه مبتنی بر مدل پگاسوس است [۱۵].

۳-۳ روش‌های مبتنی بر مدل‌های از پیش آموزش دیده

۴-۳ ایده‌های ارایه شده بهبود خلاصه سازی متون طولانی

کوتاه

از بلند

مقاله

بنویس

جیدیوتیس و همکاران شیوه‌ی تقسیم و غلبه (دنسر)^۷ را برای بهبود خلاصه سازی اسناد طولانی پیشنهاد کرده‌اند. این روش به طور خودکار خلاصه یک سند را به چند بخش تقسیم می‌کند و هر یک از این بخش‌ها را به بخش مناسب سند جفت می‌کند تا خلاصه‌های هدف متمایز ایجاد کند. شیوه‌ی معرفی شده در نظر می‌گیرد که متون طولانی به صورت بخش‌های گسسته ساختار بندی شده‌اند. برای مطابقت هر قسمت از خلاصه با بخشی از سند در دنسر از معیار روژ^۸ استفاده می‌شود. در این روش معیار روژ-ال بین هر یک از جملات خلاصه و تمام جملات سند محاسبه می‌شود و هر جمله‌ی خلاصه هدف به بخش حاوی جمله با بیشترین روژ-ال نسبت داده می‌شود. سپس تمام جملات خلاصه‌ی هدف مربوط به هر بخش را به هم الحاق می‌کنیم تا خلاصه‌ی هدف برای هر بخش ایجاد شود. در طول آموزش هر بخش از سند به همراه جمله‌ی خلاصه‌ی مربوط به آن به عنوان متن ورودی و خلاصه‌ی هدف استفاده می‌شود. مزایای این روش آموزش:

۱. تقسیم مساله به چند زیر مساله باعث کاهش پیچیدگی و ساده‌سازی مساله می‌شود.
۲. انتخاب خلاصه‌های هدف برای هر بخش بر اساس امتیازات روژ-ال هر جمله باعث تطابق بهتر و متمرکزتر بین دنباله‌های منبع و هدف ایجاد می‌شود.
۳. تقسیم هر سند آموزشی به چند جفت ورودی-هدف، نمونه‌های آموزشی بسیار بیشتری ایجاد می‌کند. این کار برای مدل‌های خلاصه‌سازی عصبی مفید است.
۴. این روش می‌تواند از مدل‌های خلاصه‌سازی مختلف از جمله شبکه‌ی عصبی بازگشتی و ترنسفورمرها استفاده کند.

هنگام کار با اسناد ساختاریافته طولانی، معمولاً همه بخش‌های سند کلیدی برای سند نیستند. اگر یک مقاله آکادمیک را به عنوان مثال در نظر بگیریم، بخش‌هایی مانند مرور ادبیات یا پیشینه در تلاش برای خلاصه کردن نکات اصلی مقاله ضروری نیستند و باعث افزودن نویز می‌شوند. بنابراین از بخش مرور ادبیات صرف نظر می‌شود و تمرکز سیستم خلاصه‌سازی فقط روی بخش‌های مقدمه، روش‌ها، نتایج و نتیجه‌گیری می‌باشد.

میتونه

حذف

بشه

⁷Divide-ANd-Conquer (DANCER)

⁸ROUGE

این مدل قابل ترکیب با پگاسوس یا مدل مولد نقطه‌ای^۹ می‌باشد. بخش کدگشا مدل مولد نقطه‌ای با ایجاد جملات تکراری مقابله می‌کند. هرچند ممکن است به خاطر تکرار اطلاعات در بخش‌های مختلف بازهم خلاصه‌ی تکراری ایجاد شود.

^۹Pointer-Generator model

فصل چهارم

روش های مبتنی بر یادگیری تقویتی

کارهای تحقیقاتی در زمینه ی یادگیری تقویتی^۱ و پردازش زبان طبیعی در سال های اخیر رشد کرده است. در یادگیری تقویتی یک عامل با محیط تعامل می کند و با آزمون و خطا، خط مشی بهینه را برای تصمیم گیری متوالی برای به حداکثر رساندن پاداش تجمعی آینده می آموزد. این پاداش می تواند یک معیار تعریف شده توسط توسعه دهنده بر اساس کار در حال حل باشد. در خلاصه سازی خودکار انتزاعی متن، نمونه هایی از چنین پاداش هایی ممکن است شامل حفظ برجستگی، مستلزم منطقی هدایت شده، و غیر افزونگی باشد.

به طور کلی، یادگیری تقویتی در سه حوزه مختلف برای بهبود خلاصه سازی خودکار استفاده می شود:

۴-۱ یادگیری تقویتی برای حل مسائل عمیق توالی به دنباله

استفاده از یادگیری تقویتی به منظور حل مسائل گوناگونی که مدل های دنباله به دنباله عمیق قادر به حل آن ها نیستند، امکانات بیشتری را فراهم می کند. به عنوان مثال، مشکلاتی مانند کمبود نوآوری در ایجاد خلاصه های خلاقانه و آموزنده و کاهش کیفیت خلاصه ها در صورت افزایش طول مقالات منبع، با استفاده از سیستم های یادگیری تقویتی و یادگیری خط مشی^۲ بهبود یافته است. علاوه بر این مدل های دنباله به دنباله عمیق را نمی توان برای خلاصه کردن طیف گسترده ای از اسناد استفاده کرد، زیرا مدلی که بر روی یک مجموعه داده آموزش داده می شود، در یک مجموعه داده دیگر به خوبی عمل نمی کند و قابلیت تعمیم ندارد. رویکردهای مبتنی بر یادگیری تقویتی می تواند این مشکل را با استفاده از گرادیان خط مشی انتقادی^۳ و ترکیب آن با یادگیری انتقالی^۴ برای انتقال دانش از یک مجموعه داده به مجموعه دیگر برطرف کنند [۸].

۴-۲ یادگیری تقویتی برای ترکیب خلاصه های استخراجی و انتزاعی

از یادگیری تقویتی برای ترکیب ویژگی های استخراجی با خلاصه انتزاعی برای استفاده از هر دو نوع خلاصه ی خودکار با الهام از رفتار انسان استفاده می شود. این مدل ها ابتدا برجسته ترین جملات را از سند ورودی استخراج می کنند، سپس با استفاده از دو شبکه: شبکه های استخراج کننده و انتزاعی، آنها را انتزاع می کنند. به عنوان مثال ليو و همکاران یک چارچوب متخاصم را پیشنهاد می کنند که مدل های انتزاعی و استخراجی را همزمان با استفاده از گرادیان خط مشی برای بهینه سازی مدل انتزاعی برای خلاصه ای با پاداش بالا، آموزش می دهد که منجر به خلاصه ای منسجم تر می شود. [۱۵]

ادیت

¹reinforcement learning

² policy learning

³self-critic policy gradient

⁴Transfer Learning (TL)

۳-۴ یادگیری تقویتی برای ایجاد معیارها و پاداش های جدید

خلاصه سازی اسناد، مانند سایر وظایف مولد زبان، اغلب به دلیل استفاده از اهداف آموزشی مبتنی بر درست‌نمایی بیشینه^۵ مورد انتقاد قرار گرفته است. درست‌نمایی بیشینه کیفیت خلاصه‌ی تولید شده را در نظر نمی‌گیرد و ممکن است خلاصه‌هایی تولید کند که فقط یک کپی از اسناد ورودی هستند، یا می‌توانند خلاصه‌هایی را بیاموزند که پر از کلمات بی‌معنی هستند. به همین دلیل، یادگیری تقویتی به عنوان جایگزینی برای بهینه‌سازی مستقیم مدل‌ها بر روی معیارهای ارزیابی و پاداش صریح به کیفیت پیش‌بینی‌های مدل استفاده شده است [۹]. معیارهای مختلفی مانند روژ-۱^۶، روژ-۲^۷، روژ-ال^۸، امتیاز اف-۱^۹ و امتیازبرت^{۱۰} به عنوان پاداش در رویکردهای یادگیری تقویتی استفاده شده است. با این حال، پارنل و همکاران استدلال می‌کنند که استفاده از امتیازات روژ به عنوان پاداش، جنبه های مهم خلاصه سازی، مانند خوانایی، روان بودن و اشتراک اطلاعات بین اسنادی در خلاصه سازی چند سندی را نادیده می‌گیرد و یک پاداش پوشش اصلاح شده همراه با یک برآوردگر گرادیان سیاست مبتنی بر اصول (ریلکس)^{۱۱} را پیشنهاد می‌دهند [۱۴، ۱]. ریلکس یک برآوردگر گرادیان خط مشی است که دارای واریانس کم و بی طرفانه است. برای مسائل یادگیری تقویتی با فضاهای کنش مداوم، مانند خلاصه سازی متن، مناسب است [۵].

در تابع ضرر بر حسب ریلکس ۴-۱ بخش اول عبارت سیاست را تشویق می‌کند تا خروجی هایی تولید کند که پاداش مورد انتظار بالایی دارند و بخش دوم سیاست را تشویق می‌کند تا خروجی هایی مشابه خروجی های تولید شده در گذشته تولید کند همچنین

r نشان دهنده‌ی پاداش $c_\phi(\tilde{z})$ یک متغیر کنترلی از پارامترهای است که انتظار می‌رود به شدت با پاداش کاهش واریانس همبستگی داشته باشد. $p(y_s)$ احتمال دنباله مشاهده شده خروجی y_s است. z دنباله نمونه های $Gumbel - Softmax$ است. \tilde{z} دنباله ای از نمونه ها از یک توزیع $Gumbel - Softmax$ مشروط بر y_s است.

$$L_{RELAX} = -[r - c_\phi(\tilde{z})] \log p(y^s) + c_\phi(z) - c_\phi(\tilde{z}) \quad (1-4)$$

⁵maximum likelihood⁶ROUGE-1⁷ROUGE-2⁸ROUGE-L⁹F1-score¹⁰BERTScore¹¹modified coverage reward along with a principled policy gradient estimator (RELAX)

فصل پنجم

نتایج

فصل ششم

جمع بندی

منابع و مراجع

- [1] Alomari, Ayham, Idris, Norisma Binti, Sabri, Aznul Qalid Md., and Alsmadi, Izzat. Deep reinforcement and transfer learning for abstractive text summarization: A review. *Comput. Speech Lang.*, 71:101276, 2022.
- [2] Andhale, Narendra and Bewoor, Laxmi A. An overview of text summarization techniques. In *2016 international conference on computing communication control and automation (ICCUBEA)*, pages 1–7. IEEE, 2016.
- [3] El-Kassas, Wafaa S., Salama, Cherif R., Rafea, Ahmed A., and Mohamed, Hoda K. Automatic text summarization: A comprehensive survey. *Expert Systems with Applications*, 165:113679, 2021.
- [4] Elman, Jeffrey L. Finding structure in time. *Cognitive science*, 14(2):179–211, 1990.
- [5] Grathwohl, Will, Choi, Dami, Wu, Yuhuai, Roeder, Geoffrey, and Duvenaud, David Kristjanson. Backpropagation through the void: Optimizing control variates for black-box gradient estimation. *ArXiv*, abs/1711.00123, 2017.
- [6] Hochreiter, Sepp and Schmidhuber, Jürgen. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- [7] Ishikawa, Kai, Ando, Shinichi, and Okumura, Akitoshi. Hybrid text summarization method based on the tf method and the lead method. In *NTCIR Conference on Evaluation of Information Access Technologies*, 2001.

- [8] Keneshloo, Yaser, Ramakrishnan, Naren, and Reddy, Chandan K. Deep transfer reinforcement learning for text summarization. ArXiv, abs/1810.06667, 2018.
- [9] Lee, Chang-Shing, Jian, Zhi-Wei, and Huang, Lin-Kai. A fuzzy ontology and its application to news summarization. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 35(5):859–880, 2005.
- [10] Liang, Yunlong, Meng, Fandong, Zhou, Chulun, Xu, Jinan, Chen, Yufeng, Su, Jinsong, and Zhou, Jie. A variational hierarchical model for neural cross-lingual summarization. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 2088–2099, Dublin, Ireland, May 2022. Association for Computational Linguistics.
- [11] Luhn, Hans Peter. The automatic creation of literature abstracts. IBM Journal of research and development, 2(2):159–165, 1958.
- [12] Malliaros, Fragkiskos D. and Skianis, Konstantinos. Graph-based term weighting for text categorization. In Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015, ASONAM '15, page 1473–1479, New York, NY, USA, 2015. Association for Computing Machinery.
- [13] Moratanch, N. and Chitrakala, S. A survey on abstractive text summarization. In 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), pages 1–7, 2016.
- [14] Parnell, Jacob, Unanue, Inigo Jauregi, and Piccardi, Massimo. A multi-document coverage reward for relaxed multi-document summarization. In Annual Meeting of the Association for Computational Linguistics, 2022.

- [15] Sherborne, Tom and Lapata, Mirella. Meta-learning a cross-lingual manifold for semantic parsing. *Transactions of the Association for Computational Linguistics*, 11:49–67, 2023.
- [16] Vaswani, Ashish, Shazeer, Noam, Parmar, Niki, Uszkoreit, Jakob, Jones, Llion, Gomez, Aidan N, Kaiser, Łukasz, and Polosukhin, Illia. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [17] Zhang, Jingqing, Zhao, Yao, Saleh, Mohammad, and Liu, Peter. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. In *International Conference on Machine Learning*, pages 11328–11339. PMLR, 2020.

واژه‌نامه‌ی فارسی به انگلیسی

Beel بیل	۱
پ	
Episodic پاسخ دوره‌ای به سوالات Question-Answering	ابر تکامل عصبی توپولوژی تقویت‌کننده Hyper Neuroevolution of Augmenting Topologies
Continues پیوسته	ازدحام ذرات Particle Swarm
ت	
Step Function تابع قدم	Natural Language استنتاج زبان طبیعی Inference
Blurring تارشدن	Evolutionary الگوریتم تکاملی Algorithm
Switch تعویض	Entropy انتروپی
Symmetry تقارن	Backpropagation انتشار به عقب
تکامل عصبی توپولوژی‌های تقویت‌کننده Neuroevolution of Augmenting Topologies	ب
Repetition تکرار	Data Driven برپایه داده
Repetition with Variation تکرار با تغییر	Support Vector . . بردار ماشین پشتیبان Machine
Multi Headed Attention . توجه چندسر	Substrate بستر
Hard Attention توجه سخت	Naive Bayes بیز ساده‌لوحانه
Soft Attention توجه نرم	

Time Series سری‌های زمانی	Sharpen تیز کردن
ش	ج
Cosine Similarity شباهت کسینوسی	Shift جابجایی
Multi Layer شبکه پرسپترون چندلایه	Population جمعیت
Perceptron	Mutation جهش
Compositional شبکه تولید الگوی ترکیبی	خ
Pattern Producing Networks	Hubber Loss خطای هوبر
Neural Network شبکه عصبی	Rooted Mean خطای ریشه مربعات
Recurrent Neural شبکه عصبی بازگشتی	Square Error
Network	د
Gated شبکه عصبی بازگشتی دروازه‌دار	Prior Knowledge دانش پیشین
Recurrent Neural Network	Gate دروازه
Feedforward Neural شبکه عصبی جلورو	Interpolation درونیابی
Network	Batch دسته
Memory Network شبکه حافظه‌ای	ر
Cold Start شروع سرد	Knowledge Tracing ردیابی دانش
ص	Copy رونوشت‌گیری
Accuracy صحت	Repeat Copy رونوشت‌گیری تکرارشونده
ف	ژ
Sharpness Factor فاکتور تیزی	Zhao ژائو
Falcon فالکن	س
ک	سر
Differentiable کامپیوتر عصبی متمایز	
Neural Computer	
Encoding کدگذاری	
Controller کنترل‌گر	

Area Under Curve ناحیه زیرمنمودار	Collier کولیر
Generation نسل	گ
Tape نوار	Gulcehre گالچره
و	Node گره
Hidden State وضعیت مخفی	Graves گریوز
ه	Discrete گسسته
Kernel هسته	ل
ی	Logits لاجیتس
Association Recall . . . یادآوری انجمنی	م
Sequence Learning . . . یادگیری ترتیبی	ماشین تورینگ عصبی . . Machine
Edge یال	ماشین تورینگ عصبی ابرتکاملی . . Evolvable Neural Turing Machine
	ماشین تورینگ عصبی پویا . . . Dynamic . . . Neural Turing Machine
	ماشین تورینگ عصبی تکاملی Evolvable . . . Neural Turing Machine
	Similarity Metric معیار شباهت
	Initialization مقداردهی اولیه
	Scalability مقیاس‌پذیری
	Attention Mechanism . . مکانیسم توجه
	Open Source منبع‌باز
	Regularity منظم‌سازی

ن

واژه‌نامه‌ی انگلیسی به فارسی

Cross-Lingual متقابل زبانی

latent representation نمایش پنهانی

یادگیری خط مشی گرادیان خط مشی انتقادی درست‌نمایی بیشینه

A شبکه تولید الگوی ترکیبی
Pattern Producing Networks

Accuracy صحت

Continues پیوسته

Area Under Curve ناحیه زیر نمودار

Controller کنترل‌گر

Association Recall یادآوری انجمنی

Copy رونوشت‌گیری

Attention Mechanism مکانیسم توجه

Cosine Similarity شباهت کسینوسی

B

D

Backpropagation انتشار به عقب

Data Driven برپایه داده

Batch دسته

Differentiable کامپیوتر عصبی متمایز
Neural Computer

Beel بیل

Discrete گسسته

Blurring تارشدن

C

Dynamic ماشین تورینگ عصبی پویا
Neural Turing Machine

Cold Start شروع سرد

Collier کولیر

E

Edge یال	Hubber Loss خطای هوبر
Episodic پاسخ دوره‌ای به سوالات Question-Answering	Hyper . . . ماشین تورینگ عصبی ابر تکاملی Evolvable Neural Turing Machine
Encoding کدگذاری	ابر تکامل عصبی توپولوژی تقویت‌کننده Hyper Neuroevolution of Augmenting
Entropy انتروپی	Topologies
Evolutionary الگوریتم تکاملی Algorithm	I
Evolvable . . . ماشین تورینگ عصبی تکاملی Neural Turing Machine	Initialization مقداردهی اولیه
F	Interpolation درونیابی
Falcon فالکن	K
Feedforward Neural . . . شبکه عصبی جلورو Network	Kernel هسته
G	Knowledge Tracing ردیابی دانش
Gate دروازه	K-Nearest k -نزدیک‌ترین همسایه Neighbours
Gated . . . شبکه عصبی بازگشتی دروازه‌دار . . Recurrent Neural Network	L
Generation نسل	Logits لاجیتس
Graves گریوز	M
Gulcehre گالچره	Memory Network شبکه حافظه‌ای
H	Multi Headed Attention . . . توجه چندسر
Hard Attention توجه سخت	Multi Layer . . . شبکه پرسپترون چندلایه . Perceptron
Head سر	Mutation جهش
Hidden State وضعیت مخفی	N
	Naive Bayes بیز ساده لوحانه
	Natural Language . . . استنتاج زبان طبیعی Inference

Neural Network شبکه عصبی	Rooted Mean . خطای ریشه مربعات خطا . Square Error
Neural Turing . . ماشین تورینگ عصبی Machine	S
تکامل عصبی توپولوژی‌های تقویت‌کننده Neuroevolution of Augmenting Topologies	Scalability مقیاس‌پذیری
Node گره	Sequence Learning . . . یادگیری ترتیبی
O	Sharpen تیزکردن
One-hot تک‌روشن	Sharpness Factor فاکتور تیزی
Open Source منبع‌باز	Shift جابجایی
P	Similarity Metric معیار شباهت
Particle Swarm ازدحام ذرات	Soft Attention توجه نرم
Population جمعیت	Substrate بستر
Prior Knowledge دانش پیشین	Support Vector . . بردار ماشین پشتیبان Machine
R	Step Function تابع قدم
Recurrent Neural . شبکه عصبی بازگشتی Network	Switch تعویض
Regularity منظم‌سازی	Symmetry تقارن
Repeat Copy . رونوشت‌گیری تکرارشونده	T
Repetition تکرار	Tape نوار
Repetition with Variation . تکرار با تغییر	Time Series سری‌های زمانی
	Z
	Zhao ژائو