



**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[[1]](#footnote-1)

**استاد راهنما:**

**دکتر حمیدرضا برادران کاشانی**

**استاد مشاور:**

**دکتر افسانه فاطمی**

**دانشجو:**

**امیررضا صدیقین**

**شهریور 1402**



**تعهدنامه اصالت اثر**

اینجانب **امیررضا صدیقین** دانشجوی مقطع کارشناسی ارشد رشته مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیک متعهد مى‌شوم که مطالب مندرج در این پایان نامه و بروندادهاى حاصل از آن، دستاورد پژوهشی اینجانب با اشراف و راهنمايىِ استاد **دکتر حمیدرضا برادران کاشانی** است و آن دسته از مطالب اين پایان نامه که حاصل تحقیقات دیگران بوده نیز طبق شيوه‌نامه‌هاى مصوّبِ ارجاع، مستند شده و در فهرست منابع و مآخذ اين پژوهش آمده است. ضمنا اظهار مى‌دارم که اين پایان نامه پيش‌تر برای دریافت هیچ مدرک تحصیلی ارائه نشده است. بديهى است دانشگاه اصفهان برای خود اين حق را قائل است که در صورت احراز و اثبات هر گونه تخلف در اين باره، مدرک تحصیلی اینجانب را از درجه اعتبار ساقط نماید و ضمن درج موضوع در جراید کثیر الانتشار، کليه امتیازات و حقوقی را که به موجب آن پس از دوران تحصیل، از آنها بهره‌مند گشته‌ام، از اینجانب سلب و به طرف ذى‌نفع مسترد کند.

|  |
| --- |
| ـ برای رساله‌هايى که با حمایت جزيىِ مراکز برون‌دانشگاهى تدوين شده است، عبارت زير تکمیل شود**:**  **اين پایان نامه در دانشگاه اصفهان و با حمایت انجام شده است.** |

|  |
| --- |
| ـ برای رساله‌هايى که طی یک قرارداد مشخص، تحت حمایت سازمان یا نهادی تدوين شده است، عبارت زير تکمیل شود**:**  **اين پایان نامه در دانشگاه اصفهان و با حمایت طی قرارداد شماره انجام شده است.** |

|  |
| --- |
| ـ برای همه رساله‌های تدوین شده در دانشگاه اصفهان عبارت زير درج شود:  **کليه حقوق مادی و معنوی مترتب بر دستاوردهای مطالعات و نوآورى‌های ناشی از پژوهش در اين پایان نامه در چارچوب آيين‌نامه مالکیت فکری و تجاری‌سازىِ دانشگاه تعيين مى‌شود.** |

**نام و نام خانوادگى دانشجو: امضاء**

**نام و نام خانوادگى استاد (/ استادان) راهنما: امضاء**

حوزه معاونت پژوهش و فناوری

**دانشگاه اصفهان**

**دانشکده مهندسی کامپیوتر**

**گروه هوش مصنوعی و رباتیک**

پايان‌نامه آقای

**امیررضا صدیقین**

دانشجوی رشته‌ی مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیک

**عنوان پایان نامه**

به عنوان بخشي از ملزومات درجه کارشناسي ارشد

در تاريخ 01/01/1402 توسط هيات داوران زير بررسي و با درجه ............ به تصويب نهايي رسيد.

1. استاد راهنمای پایان‌نامه دکتر نام استاد راهنما با مرتبه‌ی علمی مرتبه استاد راهنما از دانشگاه/موسسه نام-دانشگاه-یا-موسسه امضا
2. استاد مشاور پایان‌نامه دکتر نام استاد مشاور با مرتبه‌ی علمی مرتبه استاد مشاور از دانشگاه/موسسه نام-دانشگاه-یا-موسسه امضا
3. استاد داور داخل گروه دکتر نام داور داخلی با مرتبه‌ی علمی مرتبه علمی داور داخلی از دانشگاه/موسسه نام-دانشگاه-یا-موسسه امضا
4. استاد داور خارج از گروه دکتر نام داور خارجی با مرتبه‌ی علمی مرتبه علمی داور خارجی از دانشگاه/موسسه نام-دانشگاه-یا-موسسه امضا

مهر و امضای مدیر گروه

**سپاس ‌گزاری**

سپاس از اساتید بزرگوارم که بدون کمک آنها امکان حرکت در این مسیر وجود نداشت.

سپاس از همه افرادی که در این مسیر من را یاری نمودند.

**تقدیم به**

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

**چکیده**

متن چکیده فارسی

**کلیدواژه‌ها:** کلیدواژه‌های فارسی

# 

# فصل اول بیان مسئله‌ی پژوهشی

## مقدمه

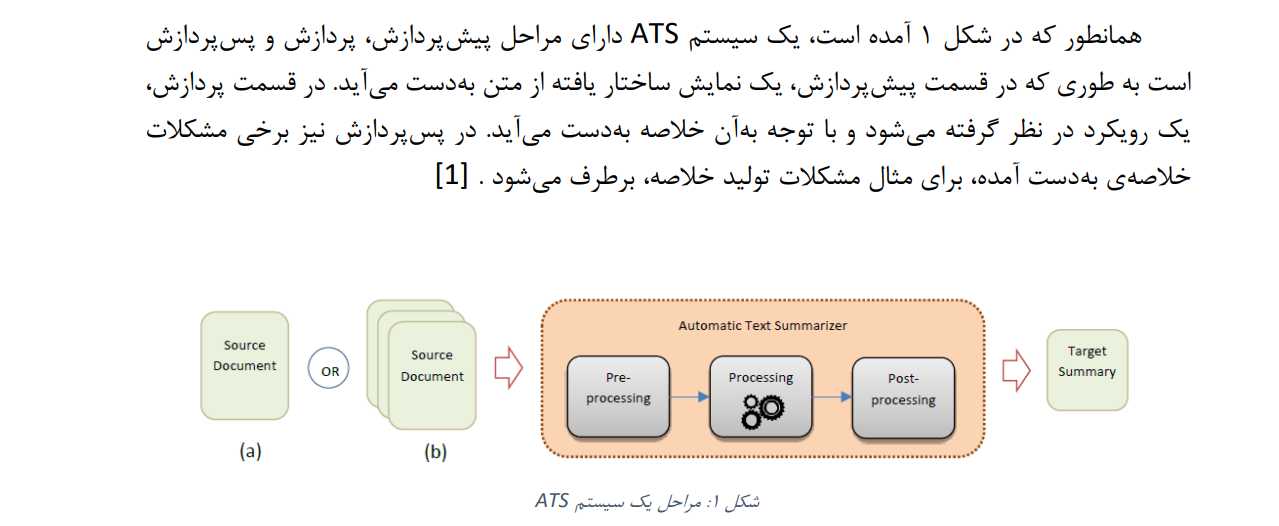
در دهه‌‌های گذشته، نحوه‌‌ی ذخیره و توزیع اطلاعات به‌ صورت چشمگیری تغییر پیدا کرده است. در حالی‌ که کتابخانه‌ها مسئولیت نگهداشت تعداد زیادی کتاب چاپی بوده است، امروزه بخش‌ زیادی از محتوای متنی به‌ صورت الکترونیکی در دسترس‌ هستند [1]. در جهان امروز، روزانه منابع متنی گسترده‌ای با سرعت بالا در حال تولید 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 2018، حاوی 6 میلیون مقاله و سند بود. پروژه‌ی Google Books بیش از 25 میلیون نسخه کتاب را تا سال 2015 دیجیتالی کرده بود. همچنین تعداد کل صفحات شاخص گذاری شده، در سپتامبر 2018 برابر 4/4 میلیارد برآورد شده است [3]. فرآیند خلاصه‌سازی دستی متن بدون شک راهی موثر برای حفظ معنا و مفهوم متن است. با این حال، خلاصه‌سازی دستی متن، یک فعالیت وقت‌گیر و پرهزینه می‌باشد [4]. به همین منظور، مسئله‌ای در حوزه‌ی پردازش زبان‌های طبیعی[[2]](#footnote-2)، به عنوان خلاصه‌سازی خودکار متن[[3]](#footnote-3) (ATS) به وجود آمده است.

خلاصه سازی 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 خلاصه سازی خودکار متن، خلاصه سازی استخراجی متن می باشد که در آن جملات مهم متن استخراج می‌شوند. اگر متن دارای ساختار منظم و جملات معینی باشد، این رویکرد می‌تواند بهترین انتخاب باشد.

شبکه‌های عصبی به علت قدرت پردازشی و توانایی فهم و تحلیل داده‌های پیچیده از اهمیت بالایی در خلاصه سازی 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 پیشرفته مانند ترنسفرمر[[4]](#footnote-4) استفاده کنند که قدرت بسیار بالایی در فهم زبان و تولید محتوا دارند [5].

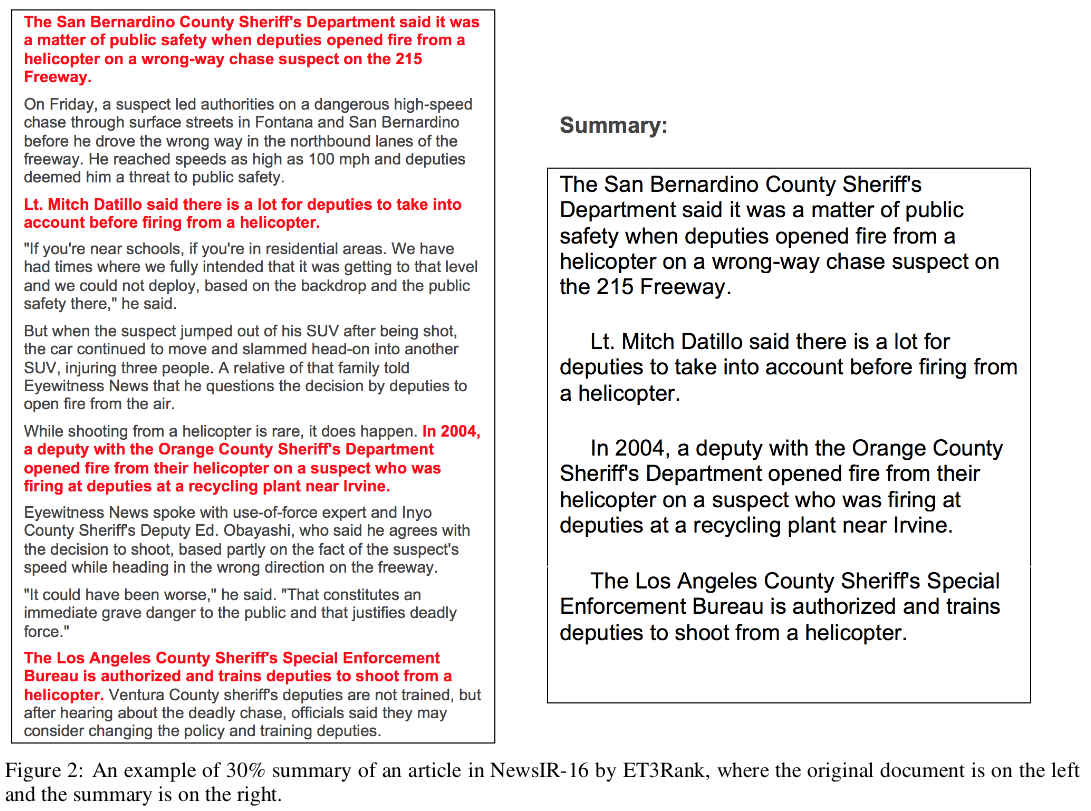
## 2-1- کلیات پژوهش

هدف سیستم‌های ATS، ساخت خلاصه‌هایی حاوی اطلاعات مهم و مرتبط با سند مربوطه، در زمان کوتاه و هزینه‌ی کم می‌باشند [3]. عمده‌ی کاربردهای سیستم‌های ATS در سیستم‌های بازیابی اطلاعات، استخراج اطلاعات و جواب‌دهی به سوالات است. برای مثال در سیستم‌های بازیابی اطلاعات، از ATS برای تقویت موتورهای جستجو استفاده می‌شود. منابع متنی می‌تواند در حوزه‌های مختلفی باشد، برای همین می‌توان کاربردهای متنوعی برای سیستم‌های ATS در نظر گرفت که خلاصه‌سازی اخبار، نظرات، کتاب، داستان، ایمیل، مقالات علمی، توییت‌ها و اطلاعات شبکه‌های اجتماعی، از موارد آن‌ها می‌باشند. همانطور که در شکل 1 آمده است، یک سیستم ATS دارای مراحل پیش‌پردازش[[5]](#footnote-5)، پردازش و پس‌پردازش[[6]](#footnote-6) 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[[7]](#footnote-7)، جداسازی کلمات توقف[[8]](#footnote-8)، هرس کردن[[9]](#footnote-9) و محاسبه‌ی ویژگی‌ها با رویکرد‌های مختلف، اشاره کرد. در قسمت پردازش، یک رویکرد خلاصه‌سازی در نظر گرفته می‌شود و با توجه به آن خلاصه به‌دست می‌آید [4]. در پس‌پردازش نیز برخی مشکلات خلاصه‌ی به‌دست آمده، برای مثال مشکل عدم تشخیص مرجع ضمایر یا مرجع زمان‌های نسبی، بر طرف می‌شود [2]. انتخاب ویژگی‌های مناسب یکی از مهم‌ترین گام‌های موجود در تولید یک مدل ATS می‌باشد. شباهت کلمات یک جمله با عنوان[[10]](#footnote-10)، موقعیت جملات در متن، ویژگی‌های TF-IDF[[11]](#footnote-11)، طول جملات، کلمات موضوعی[[12]](#footnote-12)(کلمات مرتبط با دامنه و دارای میزان تکرار زیاد)، اسامی خاص، شباهت بین جمله‌ای و اعداد، ویژگی‌های مهمی هستند که در یک سیستم ATS مورد توجه قرار می‌گیرند [4].

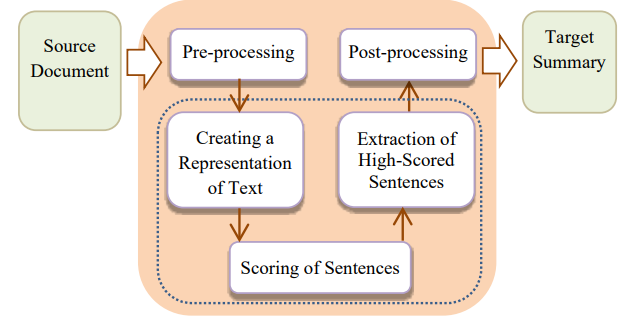


شکل 1: نمای کلی از مراحل یک سیستم ATS [4].

سیستم‌های ATS 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[[13]](#footnote-13)، انتزاعی یا اسنتتاجی[[14]](#footnote-14) و ترکیبی[[15]](#footnote-15) 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 ATS، پرداخته خواهد شد [2]. شکل ؟؟ نمونه‌ای از خلاصه‌سازی متن با رویکرد استخراجی است. همچنین شکل ؟؟ نمای کلی از مراحل یک سیستم ATS با رویکرد استخراجی را نشان می‌دهد. در این مقاله به رویکرد استخراجی در خلاصه‌سازی متن‌های تک‌سندی پرداخته شده است.



شکل 2: نمونه‌ای از خلاصه‌سازی خودکار متن با رویکرد استخراجی [6].



شکل 3: نمای کلی از معماری سیستم‌های ATS با رویکرد استخراجی [2].

برای ساخت یک سیستم ATS 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**[[16]](#footnote-16)[3]

افزونگی باعث کاهش کیفیت یک خلاصه می‌شود. وظیفة 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**[[17]](#footnote-17)[3]

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

* **مشکل پوشش ندادن تمام موضوعات اصلی**[[18]](#footnote-18)[3]

پوشش تمام موضوعات اصلی سند در خلاصه، یک موضوع اساسی در سیستم‌های ATS 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، قابل‌مشاهده هستند.

* **مشکل ناخوانایی و انسجام کم محتوا**[[19]](#footnote-19)[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 درنظرگرفتن روابط بین جمله‌ای با فاصله زیاد از هم**[[20]](#footnote-20)[7]

جملات موجود در یک سند، هرچند 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 می‌گیرند. با این‌حال مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی، نمی‌توانند به‌خوبی روابط بین جملات با فاصله‌ی زیاد از هم را در نظر بگیرند.

تاثیر شبکه‌های عصبی بر روی خلاصه سازی خودکار متن بسیار قابل توجه است. این شبکه‌ها با بهره‌گیری از معماری‌های عمیق و مدل‌های توجه، قدرتمندترین ابزارها برای انجام خلاصه سازی 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[[21]](#footnote-21)(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 خلاصه سازی خودکار متن برخوردار است [8]. همچنین تاثیر گراف و شبکه‌های عصبی گراف بر روی خلاصه سازی خودکار متن از جمله موضوعات پرطرفدار 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].

## 3-1- سوالات پژوهش

در طی پژوهش انجام شده به سوالاتی در زمینه‌ی خلاصه‌سازی خودکار متن با رویکرد استخراجی پاسخ داده شده است. در بخش نتایج و پیشینه پژوهش به میزان تاثیر شبکه‌های عصبی، به خصوص شبکه‌های عصبی گراف و میزان تاثیر ساخت گراف در مسئله‌ی ATS پرداخته شده است. همچنین این بخش‌ها، شامل مباحث دیگر از قبیل ویژگی‌ها و پیش‌پردازش‌های مناسب برای مسئله‌ی ATS می‌شوند.

## 4-1- اهداف پژوهش

با وجود روابط و متغیر‌های گوناگون در یک متن، هدف این پژوهش ارائه‌ی مدلی با رویکرد شبکه‌های عصبی جهت مدل‌سازی مناسب روابط و اطلاعات مسئله‌ی خلاصه‌سازی خودکار متن با رویکرد استخراجی می‌باشد. در این پژوهش به بررسی معماری‌های مختلف شبکه‌ی عصبی، تاثیر ساختار‌های گرافی بر روی نتایج و ویژگی‌های مناسب برای این مسئله پرداخته شده است.

## 5-1- روش ارائه شده

؟؟؟

## 6-1- 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، سنجه‌ها و معیار‌های مناسب در مسئله‌ی ATS 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 شبکه‌عصبی که در حوزه‌ی ATS 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 پنجم نتایج آزمایشهای روش ‌های 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

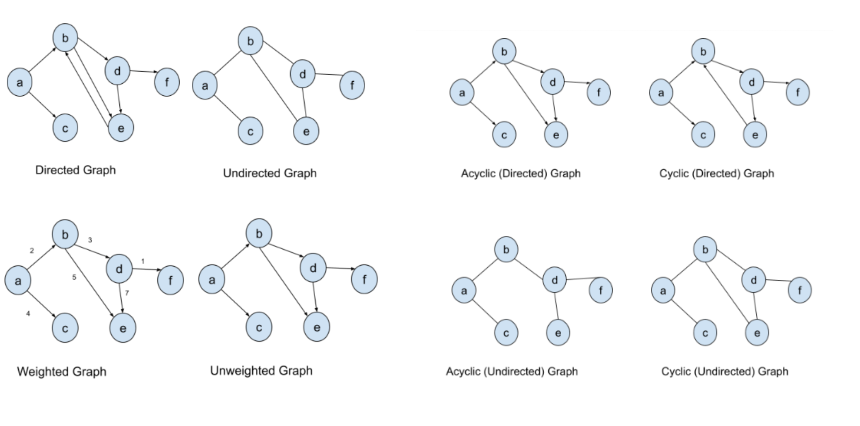
# 2-1- مقدمه

ساختار‌ها و روش‌های مختلفی برای حل مسائل خلاصه‌سازی متن ارائه شده است. در اکثر روش‌های نوین خلاصه‌سازی متن، یادگیری عمیق و ساختار‌های آن مهم‌ترین چهارچوب مورد استفاده است. از جمله شبکه‌های عصبی پیچشی (CNN)[[22]](#footnote-22) و شبکه‌های عصبی بازگشتی (RNN)[[23]](#footnote-23) از جمله شبکه‌های عصبی پرکاربرد در مسائل پردازش زبان طبیعی[[24]](#footnote-24) و پردازش متن می‌باشند. همچنین در سال‌های اخیر شبکه‌های عصبی گراف[[25]](#footnote-25) تحولات زیادی در یادگیری عمیق[[26]](#footnote-26) به وجود آورده است.

در این فصل در مورد انواع شبکه‌های عصبی و روش‌ محاسبات آن‌ها و خاصیت هر کدام توضیحاتی داده شده است. همچنین معیار‌های ارزیابی مدل‌ها در خلاصه‌سازی متن معرفی می‌گردد و در ادامه به چگونگی استفاده از شبکه‌های عصبی در پردازش زبان طبیعی به خصوص خلاصه‌سازی خودکار متن پرداخته می‌شود.

# 2-2- گراف

در ریاضیات، گراف را می‌توان به عنوان یک جفت مرتب (V,E) تعریف کرد، که در آن V مجموعه‌ای از گره‌ها[[27]](#footnote-27) (نقاط) و E مجموعه‌ای از یال‌ها[[28]](#footnote-28) (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 [11]. 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[[29]](#footnote-29)، شبکه‌های اجتماعی[[30]](#footnote-30)، پردازش زبان طبیعی، مسائل مسیریابی[[31]](#footnote-31)، و تجزیه و تحلیل شبکه‌های پیچیده[[32]](#footnote-32) کاربرد دارند.[12] همانطور که در شکل ؟؟؟ نشان داده شده است،گراف می‌تواند جهت‌دار[[33]](#footnote-33) یا بی‌جهت[[34]](#footnote-34) ، وزن‌دار[[35]](#footnote-35) یا بی وزن[[36]](#footnote-36)، گراف دوره دار[[37]](#footnote-37) و بدون دور[[38]](#footnote-38) باشد و تبع آن، مفاهیم متنوعی از گراف‌ها ایجاد می‌شود [11]. همچنین اگر گره‌های یک گراف ماهیت یکسان داشته باشند، به آن گراف همگن[[39]](#footnote-39) و اگر ماهیت گره‌ها یا هم یکسان نباشند به آن گراف ناهمگن[[40]](#footnote-40) گفته می‌شود.



شکل 4:گراف‌ها می‌توانند وزن‌دار یا بی وزن، جهت‌دار یا بی‌جهت، دوردار یا بی‌دور باشند[11].

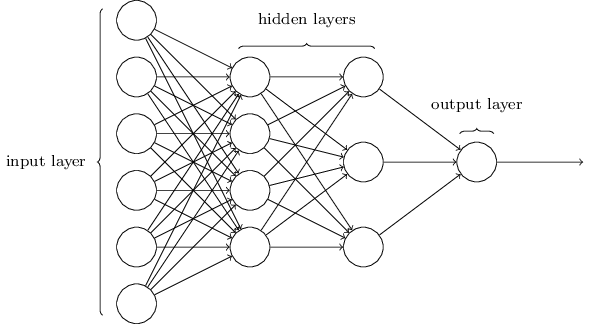
# 3-2- یادگیری ماشین[[41]](#footnote-41)

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 نظارت شده، 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 نظارت شده، 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. یادگیری ماشین، تحول عظیمی در هوش مصنوعی به وجود آورد و توانست دست‌آورد‌های زیادی را کسب کند [13], [14].

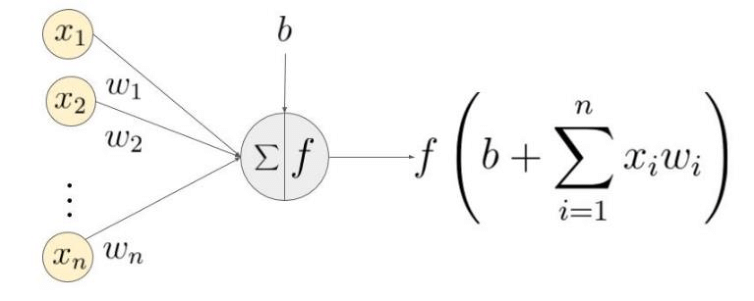
# 4-2- یادگیری عمیق

یادگیری عمیق یک حوزه پویا و حیاتی در علم داده[[42]](#footnote-42) 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 صورت و حتی بازیابی اطلاعات، به کار گرفته می‌شوند [14].

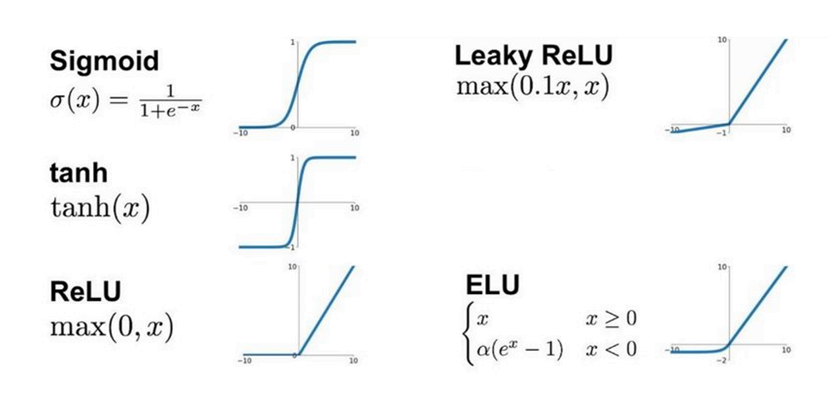
شبکه عصبی چندلایه [[43]](#footnote-43) یکی از مدل‌های اولیه و مهم در زمینه یادگیری عمیق است. این شبکه به عنوان یک نمونه از شبکه‌های عصبی عمیق، از چند لایه از نورون‌ها (واحدهای پردازشگر) تشکیل شده است. همانطور که در شکل ؟؟ آمده است، در این نوع شبکه‌ها، از سه بخش اصلی به نام لایه‌ی ورودی[[44]](#footnote-44)، لایه‌های پنهان[[45]](#footnote-45)، و لایه خروجی[[46]](#footnote-46) تشکیل شده است. در لایه ورودی، ویژگی‌های ورودی به شبکه وارد می‌شوند. لایه‌های پنهان با انجام عملیات‌های خطی و توابع غیرخطی[[47]](#footnote-47) یا فعالسازی[[48]](#footnote-48) بر 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 متصل [[49]](#footnote-49)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 خطی[[50]](#footnote-50) تفکیک پذیر نیستند، به همین منظور در برخی از لایه‌ها از توابع غیرخطی یا فعالسازی استفاده می‌کنند. همانطور که در شکل ؟؟ نحوه‌ی محاسبات درون هر نرون به نمایش گذاشته شده، ورودی هر نرون 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 نرون یک ورودی با مقدار 1 به همراه وزن آن (بایاس) نیز اضافه می‌شود. در آخر تابع فعالسازی برروی مجموع ضرب ورودی با وزن‌های متناظر اعمال می‌شود. در شکل ؟؟، برخی توابع فعالسازی متداول، معرفی شده‌اند.



شکل 5: ساختار یک شبکه عصبی که دارای سه بخش اصلی لایه‌ی ورودی، لایه مخفی و لایه‌ی خروجی میباشد [15].



شکل 6: نحوه‌ی محاسبه‌ی خروجی هر نرون در شبکه‌ها عصبی. هر نرون از چندین ورودی و وزن به همراه بایاس(مقدار وزن با ورودی 1) تشکیل شده و با استفاده از تابع غیر خطی f ویژگی غیر خطی به خروجح میدهد [16].



شکل 7: برخی توابع فعالسازی معروف[17]

متناسب با نوع مسئله و الگوریتم‌های بهینه‌سازی، فرایند آموزش شبکه‌ی عصبی شامل دو عنصر مهم تابع خطا[[51]](#footnote-51) و الگوریتم بهینه‌سازی[[52]](#footnote-52) 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 معمولا 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 برسند [14]. از الگوریتم‌های بهینه‌سازی معروف می‌توان به الگوریتم‌های SGD، Momentum، Adagrad، Adadelta، RSMProp و Adam اشاره کرد [18]. همچنین جدول ؟؟ برخی توابع خطا را معرفی می‌کند.

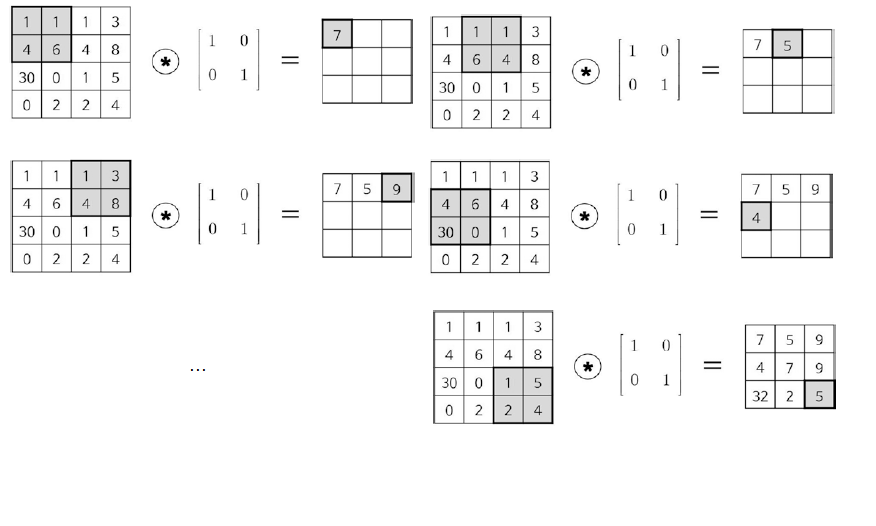
|  |  |  |
| --- | --- | --- |
| توضیحات | رابطه‌ی ریاضی | نام تابع |
| میانگین مجموع مجذور تفاضل مقدار واقعی با مقدار پیش‌بینی شده در مسائل رگرسیون. |  | MSE  (mean squared error) |
| میانگین مجموع قدرمطلق تفاضل مقدار واقعی با مقدار پیش‌بینی شده در مسائل رگرسیون. |  | MAE  (mean absolute error) |
| از توابع خطا مرسوم برای مسائل دسته‌بندی دوتایی[[53]](#footnote-53) . |  | Binary Cross Entropy |
| بسط یافته‌ی تابع binary cross entropy است و خطا را برای k کلاس محاسبه می‌کند. |  | Categorical Cross Entropy |

جدول 1: چند تابع خطا معروف در شبکه‌های عصبی[19]

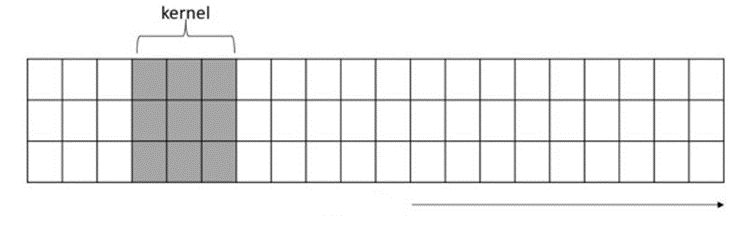
در ادامه به یک سری از شبکه‌های عصبی متداول و روش‌های معروف در این حوزه، پرداخته می‌شود.

## 1-4-2- شبکه‌های عصبی پیچشی[[54]](#footnote-54)

شبکه‌های عصبی پیچشی یا CNNها، نوع دیگری از شبکه‌های عصبی هستند که معمولا برای استخراج ویژگی‌ها و استخراج الگو‌ها استفاده می‌شود. این شبکه‌ها معمولا برای پردازش تصاویر و پردازش داده‌های توالی استفاده می‌شوند. شبکه‌های عصبی پیچشی از لایه‌های پیچشی[[55]](#footnote-55) و لایه‌های ادغام[[56]](#footnote-56) تشکیل شده‌اند. لایه‌های پیچشی به عنوان فیلترها یا کرنل‌ها عمل کرده و با حرکت این فیلترها روی داده‌ها، ویژگی‌های مختلف با محاسبه‌ی عملیات پیچشی[[57]](#footnote-57) از آن استخراج می‌کنند. عملیات پیچشی، حاصل جمع ضرب هر مؤلفه کرنل با مؤلفه نظیرش در داده می‌باشد. لایه‌های ادغام به وسیله‌ی روش‌های تجمیع[[58]](#footnote-58)، با کاهش ابعاد داده و حذف اطلاعات غیرضروری، کمک به کاهش تعداد پارامترها و افزایش سرعت آموزش می‌کنند[14] [15]. 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 ؟؟، 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 ؟؟ نحوه‌ی حرکت کرنل پیچشی یک بعدی را نشان می‌دهد. در فرایند آموزش این شبکه‌ها، مؤلفه‌های شبکه بروزرسانی شده و الگو‌های مناسب از داده‌ها استخراج می‌شود.



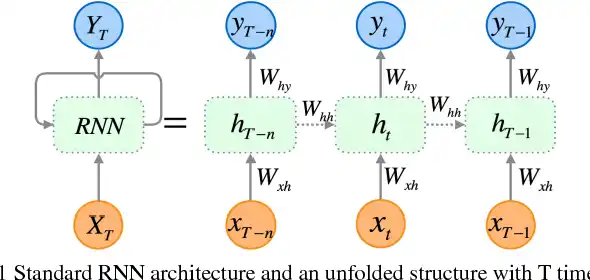
شکل 8 : مراحل بدست آوردن خروجی از اعمال عملیات پیچش یک فیلتر بر روی یک ماتریس ورودی [22] .



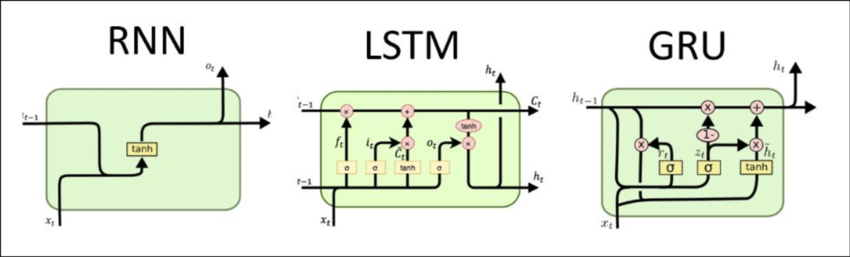
شکل 9: نحوه‌ی حرکت کرنل پیچشی یک بعدی[23]

## 2-4-2- شبکه‌های عصبی بازگشتی [[59]](#footnote-59)

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 ؟؟ نمایش داده شده است، این نوع از شبکه‌های عصبی از ویژگی حافظه‌ی بازگشتی[[60]](#footnote-60) 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[[61]](#footnote-61)، تشخیص گفتار[[62]](#footnote-62)، پیش‌بینی متن[[63]](#footnote-63)، و حتی تولید متون[[64]](#footnote-64) 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[[65]](#footnote-65) و GRU‌[[66]](#footnote-66) که در به نمایش درآمده‌اند، از متداول‌ترین شبکه‌های عصبی بازگشتی می‌باشند که در زمینه‌ی این چالش‌ها بهینه شده‌اند [24]. شکل ؟؟، معماری‌های GRU و LSTM و تفاوت آن‌ها با RNNهای اولیه را نمایش می‌دهد.



شکل 10: نحوه‌ی عملکرد شبکه‌های عصبی بازگشتی[25] .

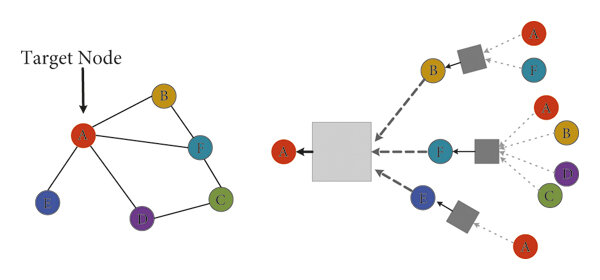


شکل 11: شبکه‌های GRU و LSTM و تفاوت آن‌ها با شبکه‌های RNN اولیه [26].

## 3-4-2- شبکه‌های عصبی گراف[[67]](#footnote-67)

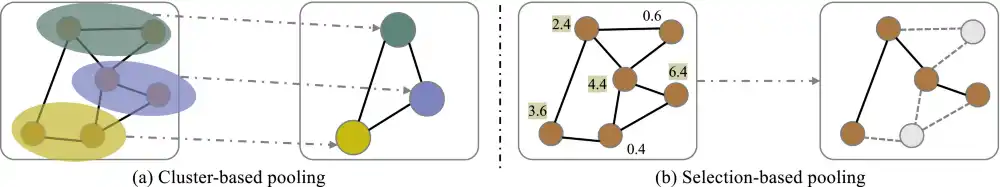
شبکه‌های عصبی گراف یک دسته از مدل‌های یادگیری عمیق هستند که برای مدل‌سازی داده‌های گرافی مانند شبکه‌های اجتماعی، شبکه‌های مولکولی و ساختارهای گرافی مشابه به کار می‌روند. این شبکه‌ها توانمندی بالایی در تفسیر و تحلیل روابط پیچیده و الگوهای شبکه‌های گرافی دارند. به عبارت دیگر، آنها قابلیت مدل‌سازی تعاملات و وابستگی‌های بین گره‌ها و لینک‌های یک گراف را دارا هستند [27].

شبکه‌های عصبی گراف از مکانیزم انتقال پیام[[68]](#footnote-68)، 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 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[[69]](#footnote-69) نیز در نظر گرفت. مدیریت پیچیدگی مدل‌ها، تعامل با ساختار گراف، انعطاف پذیری در مقابل اندازه‌ی گراف‌ها، آموزش ناهمگن[[70]](#footnote-70) گراف‌ها و مقاومت در برابر نویز از قابلیت‌های شبکه‌های عصبی می‌باشد[27]. یکی از اولین و متداول‌ترین شبکه‌های عصبی گراف، GCN‌ها[[71]](#footnote-71) می‌باشند که در شکل ؟؟، نحوه‌ی عملکرد آن نمایش داده شده است. یکی دیگر از شبکه‌های عصبی گراف معروف، شبکه‌ی عصبی GAT‌[[72]](#footnote-72) می‌باشد که با استفاده از مکانیزم توجه، می‌تواند میزان اهمیت بین گره‌ها در فرایند انتقال پیام را آموزش ببیند.



شکل 12: نحوه‌ی انتقال پیام برای یک گره مشخص در شبکه‌ی GCN [28]

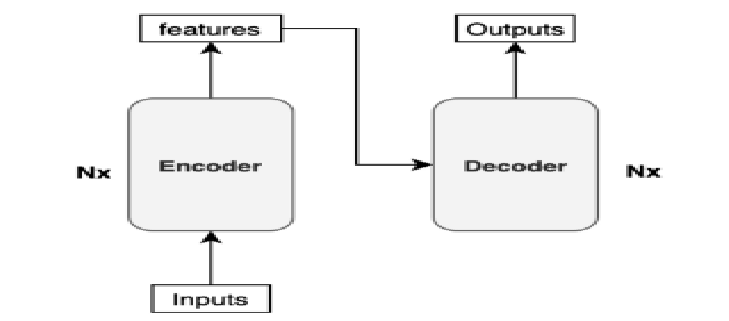
برخی مواقع در شبکه‌های عصبی گراف از لایه‌های ادغامی گراف[[73]](#footnote-73) به منظور به‌دست آوردن گراف کوچک‌تر استفاده می‌شود. همچنین با این روش می‌توان نمایشی برای گراف مورد نظر نیز به‌دست آورد. در شکل ؟؟ یک نمونه از ادغام گره‌های گراف نمایش داده شده است.



شکل 13: ادغام گره‌های گراف و تولید گراف کوچکتر با استفاده از لایه‌ی ادغام گراف [29]

# 5-2- ساختار کدگذار-کدگشا [[74]](#footnote-74)

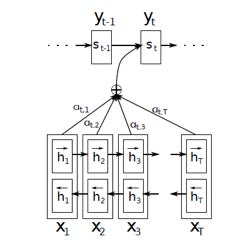
ساختار‌های مختلفی برای حل مسائل متفاوت بر بستر یادگیری عمیق ارائه شده‌اند. یکی از این ساختار‌ها، ساختار کدگذار-کدگشا است. ساختار کدگذار-کدگشا یک معماری متداول در حوزه شبکه‌های عصبی است که برای مسائل مختلفی از جمله ترجمه ماشینی، تولید متن، و تولید تصویر[[75]](#footnote-75) مورد استفاده قرار می‌گیرد. کدگذار مسئول تبدیل ورودی به یک فضانی نهان[[76]](#footnote-76) یا بردار ویژگی است. 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 ؟؟، ساختار کلی کدگذار-کدگشا را نشان می‌دهد.



شکل 14: ساختار کدگذار-کدگشا [31]

# 6-2- روش‌های مبتنی بر توجه[[77]](#footnote-77)

یکی از ساختار‌های مورد استفاده در سال‌های گذشته، ساختار توجه 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 کمک کند. این مفهوم ابتدا در سال 2015، در معماری ترنسفرمر[[78]](#footnote-78) معرفی شد و از آن زمان به عنوان یک عنصر اساسی در بسیاری از مدل‌های پیشرفته عصبی مورد استفاده قرار گرفته است [5]. شکل ؟؟ ساختار توجه در فرایند بدست آوردن نمایش بهتر برای دنباله را نشان می‌دهد.



شکل 15: نمایش مکانیزم توجه برای بدست آوردن نمایش برای یک دنباله[33]

در هر سلول توجه[[79]](#footnote-79)، سه مفهوم کلید[[80]](#footnote-80)، مقدار[[81]](#footnote-81) و پرسش[[82]](#footnote-82) 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 (؟؟) محاسبه می‌شود [5].

در عبارات ریاضی بالا منظور از q پرسش، k کلید و v مقدار می‌باشد. تابع SIM، یک تابع شباهت می‌باشد. معادله‌ی 2 نیز نحوه‌ی محاسبه‌ی SoftMax است که در آن ضریب اهمیت، محاسبه می‌شود. یکی از روش‌های محاسبه‌ی شباهت می‌تواند ضرب داخلی مقیاس‌بندی شده[[83]](#footnote-83) باشد. معادله (؟؟) نحوه‌ی محاسبه‌‌ی ضرب داخلی مقیاس‌ بندی شده را نشان می‌دهد [5].

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[[84]](#footnote-84) 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[[85]](#footnote-85) گویند [34].

# 7-2- جانشانی کلمه[[86]](#footnote-86)

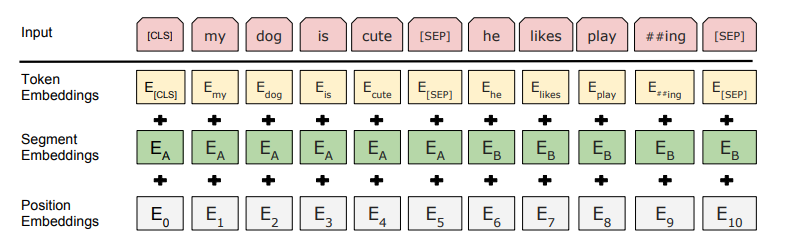
در پردازش زبان طبیعی، برای استفاده از روش‌های یادگیری ماشین، کلمات یا عبارات باید به بردار‌های عددی تبدیل شوند. بنابراین روشی برای تبدیل متن به بردار نیاز است. جانشانی کلمه یک روش در زمینه پردازش زبان طبیعی است که اهمیت زبانی کلمات را در یک فضای چندبعدی عددی نمایش می‌دهد. این تکنیک مفهومی از معانی و 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 معنای 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 [35]. از روش‌های معروف در این حوزه می‌توان به word2vec، glove‌[[87]](#footnote-87) و برخی مدل‌های از پیش آموزش دیده شده[[88]](#footnote-88) مثل مدل BERT‌[[89]](#footnote-89) نام برد.

## 1-7-2- 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 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 [36]. Glove به عنوان یکی از مدل‌های موفق در زمینه جانشانی کلمه به شناخته شده است و به طور گسترده در پروژه‌های پردازش زبان طبیعی مورد استفاده قرار گرفته است.

## 3-7-2- 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[[90]](#footnote-90)، 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[[91]](#footnote-91) 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[[92]](#footnote-92)، استخراج ارتباطات[[93]](#footnote-93)، ترجمه ماشینی، و حتی فهم متن وابسته به سوالات[[94]](#footnote-94) استفاده شوند. همانطور که در شکل ؟؟ آمده است، این مدل از سه جانشانی توکن، بخش و موقعیت، برای بدست آوردن جانشانی کلمه استفاده می‌کند. این معماری و آموزش آن بر روی مقدار زیادی داده باعث شده است تا جانشانی کلمات با استفاده از ویژگی‌های زمینه‌ای[[95]](#footnote-95) استخراج شوند [37].



شکل 16: جانشانی کلمه در مدل BERT با استفاده از سه جانشانی توکی، بخش و موقعیت تولید می‌شود.

# 8-2- 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 پیداکردن نمایش عددی، انجام شده است. معیار‌های ارزیابی یک سیستم خلاصه‌سازی خودکار متن به دو دسته‌ی ذاتی[[96]](#footnote-96) و بیرونی[[97]](#footnote-97) تقسیم می‌شوند. معیار‌های بیرونی میزان پیروی از یک ساختار (بیشتر مواردی مثل مسائل پاسخ‌دهی به سؤالات[[98]](#footnote-98) یا بازیابی اطلاعات[[99]](#footnote-99)) را نمایش می‌دهند. اما معیار‌های ذاتی بر روی انسجام و میزان کیفیت معنایی، تمرکز دارند. معیار‌های ذاتی نیز خود به دو دسته‌ی، معیار‌های ارزیابی کیفیت متن[[100]](#footnote-100) و ارزیابی محتوا[[101]](#footnote-101) تقسیم می‌شوند. در سیستم‌های خلاصه‌سازی متن، معمولاً از معیار‌های ذاتی استفاده می‌شود [38], [39]. در ادامه به برخی از معروف‌ترین و پرکاربردترین معیار‌های ارزیابی در این سیستم‌ها اشاره خواهد شد.

## 1-8-2- معیار صحت یا **Precision**

معیار صحت یا precision تعیین می‌کند که خلاصه‌ی به‌دست‌آمده، به چه میزان مختصر و فاقد کلمات اضافی می‌باشد. میزان این معیار مطابق با رابطه‌ی ؟؟ بدست می‌آید [39].

دراین‌رابطه، نسبت تعداد کلمات مشترک خلاصه‌ی مرجع () و خلاصه‌‌ی حاصل از سیستم ATS () به تعداد کل کلمات خلاصه‌ی بدست آمده، محاسبه می‌شود .

## 2-8-2- معیار **Recall**

معیار Recall میزان پوشش خلاصه‌ی مرجع توسط خلاصه‌ی حاصل شده را تعیین می‌کند. مقدار این معیار با استفاده از رابطه‌ی ؟؟ بدست می‌آید [39].

## 3-8-2- معیار **F-measure**

به منظور ادغام دو معیار Recall و precision، از معیار F-measure استفاده می‌شود. رابطه‌ی ؟؟ مقدار این معیار را به دست می‌آورد [39].

## 3-8-2- معیار**ROUGE‌‌[[102]](#footnote-102)**

معیار ROUGE مرسوم‌ترین معیار ارزیابی در سیستم‌های خلاصه‌سازی خودکار متن است. معیار ROUGE با استفاده از n-gram‌های مشترک بین خلاصه‌ی مرجع و خلاصه‌ی تولید شده، یک مقدار عددی برای کیفیت محتوای خلاصه ارائه می‌دهد. رابطه‌ی ؟؟ میزان عددی ROUGE متناسب با n-gram موردنظر را نشان می‌دهد [40].

رابطه‌ی فوق مقدار Rouge با رویکرد Recall را نشان می‌دهد که نسبت مجموع n-gramهای مشترک در جملات خلاصه‌ی مرجع و خلاصه‌ی تولید شده () به مجموع n-gramهای مشترک در جملات خلاصه‌ی مرجع () را بدست می‌آورد [39]. همانطور که در رابطه‌ی ؟؟ آمده است، Rouge با رویکرد Perecision محاسبه می‌شود و با توجه به دو معیار Perecision و Recall می‌توان Rouge با معیار F-measure را به صورت معادله‌ی ؟؟ محاسبه کرد.

در مسئله‌ی خلاصه‌سازی خودکار متن 3 معیار R1، R2 و RL مرسوم‌ترین شاخص‌های گزارش شده در مقالات می‌باشند. R1 به معنای محاسبه‌ی ROUGE با n برابر 1 می‌باشد. این شاخص میزان همپوشانی توکن‌های یکتایی[[103]](#footnote-103) بین خلاصه‌ی سیستم و خلاصه‌ی مرجع را نشان می‌دهد. همچنین R2 به معنای ROUGE با n برابر 2 است که نشان‌دهنده‌ی همپوشانی توکن‌های دوتایی[[104]](#footnote-104) بین خلاصه‌ی سیستم و خلاصه‌ی مرجع می‌باشد. RL نیز میزان ROUGE بر اساس طول بزرگترین زیردنباله‌ی مشترک بین دو خلاصه را نشان می‌دهد [39]. 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 3 کلمه در نظر گرفته می‌شوند.

# 9-2- جمع‌بندی

در این فصل به بررسی مفاهیم‌ مورد استفاده در این پژوهش پرداخته شد. ابتدا برخی روابط ریاضی و انواع گراف‌ها معرفی شدند. سپس یادگیری ماشین و یادگیری عمیق و برخی تکنیک‌های آن ها معرفی گردید. یکی از بزرگ‌ترین مزیت‌های یادگیری عمیق استخراج ویژگی‌های نهان داده‌های ورودی می‌باشد. شبکه‌های عصبی معمولا 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 word2vec، glove‌ 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 معمولا در پژوهش‌های مربوط به حوزه‌ی خلاصه‌سازی متن گزارش می‌شوند.

# فصل سوم پیشینه پژوهش

# 1-3- مقدمه

رویکرد استخراجی خلاصه‌سازی متن را می‌توان به‌صورت یک مسئله‌ی دسته‌بندی جملات[[105]](#footnote-105) در نظر گرفت، بدین صورت که باید مدل تشخیص دهد، هر جمله در دسته‌ی خلاصه قرار بگیرد یا خیر[41]. پس هر مدل استخراجی، از یک دسته‌بند[[106]](#footnote-106) برای عملیات دسته‌بندی جملات، استفاده می‌کند. این دسته‌بند‌ها می‌تواند به صورت یک شبکه‌ی عصبی (MLP، LSTM و ...) یا به صورت الگوریتم‌هایی مثل الگوریتم‌های رتبه‌بندی[[107]](#footnote-107) 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 [42]. 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 شبکه‌های عصبی در زمینه‌ی خلاصه‌سازی بوده است، که توانسته است پیشرفت قابل توجهی بدست آورد [41], [43]. 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[[108]](#footnote-108)، بهبود‌های بسیاری در این زمینه رخ داد [41]. 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 [44], [45]. شبکه‌های عصبی گراف به چهار دسته‌ی شبکه‌های عصبی گراف بازگشتی[[109]](#footnote-109)، شبکه‌های پیچشی گراف[[110]](#footnote-110)، خود رمزگذار گراف[[111]](#footnote-111) و شبکه‌های عصبی گراف فضایی-زمانی[[112]](#footnote-112) تقسیم می‌شوند [46]. 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 [47].

روش‌های موجود در خلاصه‌سازی استخراجی متن به دو دسته‌ی کلی تقسیم می‌شوند: 1- رویکردهای کلاسیک و 2- رویکردهای مبتنی بر ماشین لرنینگ.

رویکردهای کلاسیک 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.

بر اساس ادعاهای موجود در مطالعات، مدل‌های مبتنی بر ماشین لرنینگ اغلب دقت و کارآیی بهتری در خلاصه‌سازی استخراجی نسبت به رویکردهای سنتی[[113]](#footnote-113) ارائه کرده‌اند [43]. 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.

# 2-3- 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-IDF‌[[114]](#footnote-114)) 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 [48]. 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 به‌دست آورد [49].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 کلمات عنوان بیشتر باشد، جمله به میزان بیشتری مرتبط خواهد بود [48]. 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.

Top of Form

1. Extractive text summarization using deep learning approaches [↑](#footnote-ref-1)
2. Natural language processing [↑](#footnote-ref-2)
3. Automatic text summarization [↑](#footnote-ref-3)
4. Transformer [↑](#footnote-ref-4)
5. Preprocessing [↑](#footnote-ref-5)
6. Postprocessing [↑](#footnote-ref-6)
7. Part of speech tagging [↑](#footnote-ref-7)
8. Stop words Filtering [↑](#footnote-ref-8)
9. Stemming [↑](#footnote-ref-9)
10. Title similarity [↑](#footnote-ref-10)
11. Term frequency – inverse term frequency [↑](#footnote-ref-11)
12. Thematic words [↑](#footnote-ref-12)
13. Extractive [↑](#footnote-ref-13)
14. Abstractive [↑](#footnote-ref-14)
15. Hybrid [↑](#footnote-ref-15)
16. Problem of redundancy [↑](#footnote-ref-16)
17. Problem of irrelevancy [↑](#footnote-ref-17)
18. Problem of loss of coverage [↑](#footnote-ref-18)
19. Problem of non-readability and less cohesive content [↑](#footnote-ref-19)
20. Long-distance relationships [↑](#footnote-ref-20)
21. Recurrent Neural Network [↑](#footnote-ref-21)
22. Convolutional neural networks [↑](#footnote-ref-22)
23. Recurrent neural networks [↑](#footnote-ref-23)
24. Natural language processing (NLP) [↑](#footnote-ref-24)
25. Graph neural networks [↑](#footnote-ref-25)
26. Deep learning [↑](#footnote-ref-26)
27. Nodes [↑](#footnote-ref-27)
28. Edges [↑](#footnote-ref-28)
29. Data architecture [↑](#footnote-ref-29)
30. Social networks [↑](#footnote-ref-30)
31. Routing problems [↑](#footnote-ref-31)
32. Analysis of complex networks [↑](#footnote-ref-32)
33. Directed graph [↑](#footnote-ref-33)
34. Undirected graph [↑](#footnote-ref-34)
35. Weighted graph [↑](#footnote-ref-35)
36. Unweighted graph [↑](#footnote-ref-36)
37. Cyclic graph [↑](#footnote-ref-37)
38. Acyclic graph [↑](#footnote-ref-38)
39. Homogeneous graph [↑](#footnote-ref-39)
40. Heterogeneous graph [↑](#footnote-ref-40)
41. Machine learning [↑](#footnote-ref-41)
42. Data science [↑](#footnote-ref-42)
43. Multi-layer perceptron (MLP) [↑](#footnote-ref-43)
44. Input layer [↑](#footnote-ref-44)
45. Hidden layers [↑](#footnote-ref-45)
46. Output layer [↑](#footnote-ref-46)
47. Non-linear [↑](#footnote-ref-47)
48. Activation functions [↑](#footnote-ref-48)
49. Fully connected [↑](#footnote-ref-49)
50. Linear [↑](#footnote-ref-50)
51. Loss function [↑](#footnote-ref-51)
52. Optimizer algorithms [↑](#footnote-ref-52)
53. Binary classification [↑](#footnote-ref-53)
54. Convolutional neural network [↑](#footnote-ref-54)
55. Conolutional layers [↑](#footnote-ref-55)
56. Pooling layers [↑](#footnote-ref-56)
57. Convolution [↑](#footnote-ref-57)
58. Aggregation methods [↑](#footnote-ref-58)
59. Recurrent neural networks [↑](#footnote-ref-59)
60. Recurrent memory [↑](#footnote-ref-60)
61. Machine translation [↑](#footnote-ref-61)
62. Speech recognition [↑](#footnote-ref-62)
63. Text prediction [↑](#footnote-ref-63)
64. Text generation [↑](#footnote-ref-64)
65. Long short-term memory [↑](#footnote-ref-65)
66. Gated recurrent unit [↑](#footnote-ref-66)
67. Graph neural networks [↑](#footnote-ref-67)
68. message passing [↑](#footnote-ref-68)
69. relations [↑](#footnote-ref-69)
70. Heterogeneous Learning [↑](#footnote-ref-70)
71. Convolutional network [↑](#footnote-ref-71)
72. Graph attention networks [↑](#footnote-ref-72)
73. Graph pooling layers [↑](#footnote-ref-73)
74. Encoder-decoder structure [↑](#footnote-ref-74)
75. Image generation [↑](#footnote-ref-75)
76. Latent space [↑](#footnote-ref-76)
77. Attention-based methods [↑](#footnote-ref-77)
78. transformer [↑](#footnote-ref-78)
79. Attention cell [↑](#footnote-ref-79)
80. key [↑](#footnote-ref-80)
81. value [↑](#footnote-ref-81)
82. query [↑](#footnote-ref-82)
83. Scaled dot-product [↑](#footnote-ref-83)
84. Multi-head attention [↑](#footnote-ref-84)
85. Self-attention [↑](#footnote-ref-85)
86. Word embedding [↑](#footnote-ref-86)
87. Global vectors for word representation [↑](#footnote-ref-87)
88. Pretrained model [↑](#footnote-ref-88)
89. Bidirectional encoder representations from transformers [↑](#footnote-ref-89)
90. bidirectional [↑](#footnote-ref-90)
91. token [↑](#footnote-ref-91)
92. Named entity recognition (NER) [↑](#footnote-ref-92)
93. Relation extraction (RE) [↑](#footnote-ref-93)
94. Question answering (QA) [↑](#footnote-ref-94)
95. Contextual feature [↑](#footnote-ref-95)
96. Intrinsic [↑](#footnote-ref-96)
97. Extrinsic [↑](#footnote-ref-97)
98. Question answering [↑](#footnote-ref-98)
99. Information retrieval [↑](#footnote-ref-99)
100. Text quality evaluation [↑](#footnote-ref-100)
101. Content evaluation [↑](#footnote-ref-101)
102. Recall-Oriented Understudy for Gisting Evaluation [↑](#footnote-ref-102)
103. Unigram [↑](#footnote-ref-103)
104. Bigram [↑](#footnote-ref-104)
105. Sentence classification [↑](#footnote-ref-105)
106. classifier [↑](#footnote-ref-106)
107. Ranking algorithms [↑](#footnote-ref-107)
108. Contextual representations [↑](#footnote-ref-108)
109. Recurrent graph neural networks [↑](#footnote-ref-109)
110. Convolutional graph neural networks [↑](#footnote-ref-110)
111. Graph AutoEncoder [↑](#footnote-ref-111)
112. spatial-temporal graph neural networks [↑](#footnote-ref-112)
113. classical approaches [↑](#footnote-ref-113)
114. Term frequency-inverse document frequency [↑](#footnote-ref-114)