دانشگاه تهران
دانشکده علوم و فنون نوین
گروه فناوری بین‌رشته‌ای

ارزیابی و مقایسه بازنمایی‌های معنایی با بهره‌گیری از علم شبکه

پایان‌نامه برای دریافت درجهٔ کارشناسی ارشد در رشتهٔ زبانشناسی رایانشی

مهنا هویدا

اساتید راهنما

دکتر مصطفی صالحی و دکتر محمود بیجن‌خان

بهمن ۱۴۰۱


بسم الله الرحمن الرحيم

دانشگاه تهران
دانشکده علوم و فنون نوین
گروه فناوری بین‌رشته‌ای

ارزیابی و مقایسه بازنمایی‌های معنایی با بهره‌گیری از علم شبکه

پایان‌نامه برای دریافت درجهٔ کارشناسی ارشد در رشتهٔ زبانشناسی رایانشی

مهنا هویدا

اساتید راهنما

دکتر مصطفی صالحی و دکتر محمود بیجن‌خان

استاد مشاور

دکتر Paulino Villas Boas

بهمن ۱۴۰۱

دانشگاه تهران

دانشکده علوم و فنون نوین

گواهی دفاع از پایان‌نامهٔ کارشناسی ارشد

هیأت داوران پایان‌نامهٔ کارشناسی ارشد آقای / خانم مهنا هویدا به شمارهٔ دانشجویی ۸۳۰۴۹۹۰۹۹ در رشتهٔ زبانشناسی رایانشی – گرایش را در تاریخ ..... با عنوان «ارزیابی و مقایسه بازنمایی‌های معنایی با بهره‌گیری از علم شبکه»

	به عدد	به حروف
با نمرهٔ نهایی		

و درجهٔ [ ] ارزیابی کرد.

ردیف	مشخصات هیأت داوران	نام و نام خانوادگی	مرتبهٔ دانشگاهی	دانشگاه یا مؤسسه	امضا
۱	استاد راهنما	دکتر مصطفی صالحی	دانشیار	دانشگاه تهران	
۲	استاد راهنما	دکتر محمود بیجن‌خان	استاد	دانشگاه تهران	
۳	استاد مشاور	دکتر Paulino Villas Boas	پژوهشگر	موسسه تحقیقاتی امبراپا	
۴	استاد داور داخلی	دکتر هادی ویسی	دانشیار	دانشگاه تهران	
۵	استاد مدعو	دکتر محمدرضا ابوالقاسمی دهاقانی	استادیار	پردیس فنی دانشگاه تهران	

نام و نام خانوادگی معاون تحصیلات تکمیلی و

پژوهشی دانشکده / گروه:

تاریخ و امضا:

تعهدنامهٔ اصالت اثر

باسمه تعالی

اینجانب مهنا هویدا تأیید می‌کنم که مطالب مندرج در این پایان‌نامه حاصل کار پژوهشی اینجانب است و به دستاوردهای پژوهشی دیگران که در این نوشته از آن‌ها استفاده شده است مطابق مقررات ارجاع گردیده است. این پایان‌نامه قبلاً برای احراز هیچ مدرک هم‌سطح یا بالاتری ارائه نشده است.

نام و نام خانوادگی دانشجو: مهنا هویدا

تاریخ و امضای دانشجو:

تقدیم به:

پدر و مادرم

قدردانی

سپاس خداوندگار حکیم را که با لطف بی‌کران خود، آدمی را به زیور عقل آراست.

در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ اساتید راهنمای خود، جناب آقای دکتر مصطفی صالحی و دکتر محمود بیجن‌خان، و استاد مشاور آقای Paulino Villas Boas صمیمانه تشکر و قدردانی کنم که در طول انجام این پایان‌نامه با نهایت صبوری همواره راهنما و مشوق من بودند و قطعاً بدون راهنمایی‌های ارزنده ایشان، این مجموعه به انجام نمی‌رسید.

و در پایان، بوسه می‌زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می‌کنم وجود مقدس‌شان را و تشکر می‌کنم از خانواده عزیزم به پاس عاطفه سرشار و گرمای امیدبخش وجودشان، که بهترین پشتیبان من بودند.

مهنا هویدا

بهمن ۱۴۰۱

چکیده

بازنمایی‌های معنایی جز مهمی از مدل‌های پردازش زبان طبیعی هستند، چرا که امکان پردازش و درک معنای عبارات و در نتیجه انجام امور زبانی گسترده‌ای را برای مدل‌های زبانی فراهم می‌کنند. با وجود اینکه تا امروز روش‌های مختلفی به منظور دست‌یابی به یک بازنمایی معنایی معرفی شده است، اطلاعات محدودی درباره ویژگی‌های ساختاری بازنمایی حاصل از این روش‌ها داریم. هم‌چنین مقایسه جامعی میان این بازنمایی‌ها صورت نگرفته است. این پژوهش به هدف مطالعه و طبقه‌بندی بازنمایی‌های معنایی براساس ویژگی‌های ساختاری گراف آن‌ها انجام شده است. به این منظور با بهره‌گیری از شاخص‌های علم شبکه بازنمایی‌های معنایی حاصل از مدل‌سازی توزیعی معنا و هم‌چنین بازنمایی‌های معنایی مبتنی بر دانش انسانی را مطالعه کرده‌ایم. این مطالعه در دو مقیاس سراسری و میانی شبکه انجام شده است. هم‌چنین، با توجه به تفاوت اندازه بازنمایی‌ها، رویکرد آماری نوینی اتخاذ کرده‌ایم که طبقه‌بندی بازنمایی‌های معنایی را بهبود می‌بخشد. با بررسی نتایج دریافتیم که در بازنمایی‌های مبتنی بر دانش انسانی کلمات پرتکرار زبان انگلیسی نسبت قابل توجهی از روابط معنایی را به خود اختصاص می‌دهند، این در حالی است که در بازنمایی‌های مبتنی بر مدل‌سازی توزیعی معنا، کلمات بسیار نادر روابط معنایی گسترده‌ای دارند. دوما مشاهده کردیم که رویکرد آماری اتخاذ شده مقایسه منسجم‌تری میان بازنمایی‌های معنایی غیریکسان فراهم می‌کند. همچنین از مقایسه بازنمایی‌های پایه با بازنمایی‌های ترکیبی دریافتیم که افزودن اطلاعات به یک بازنمایی پایه می‌تواند منجر به شکل‌گیری ویژگی های ساختاری متفاوتی شود؛ برای مثال دریافتیم که در بازنمایی‌های توزیعی که علاوه بر متن از اطلاعات تصویری نیز بهره می‌گیرند احتمال تشکیل گروه‌های معنایی کاهش می‌یابد، در حالی که این احتمال در بازنمایی‌هایی ترکیبی حاصل از دانش انسانی و مدل‌سازی توزیعی معنا، افزایش می‌یابد. بنا بر اطلاعات ما، این اولین پژوهشی است که در یک مقایسه جامع هفت بازنمایی معنایی انسانی و توزیعی را مورد مطالعه قرار داده است و روش مقایسه آماری بازنمایی‌های غیرهم‌اندازه نیز برای نخستین بار در این پژوهش ارائه شده است. در همین راستا، مسیرهای مختلف گسترش این پژوهش در آینده نیز ارائه شده است.

واژگان کلیدی بازنمایی معنایی، شبکه معنایی، مدل معنایی توزیعی، نظریه گراف، شبکه پیچیده، مقایسه شبکه‌های غیرهم‌اندازه، مدل پیکربندی تصادفی

فهرست مطالب

فهرست تصاویر

فهرست جداول

فهرست الگوریتم‌ها

فصل ۱

مقدمه

پردازش زبان طبیعی زیرشاخه‌ای از هوش مصنوعی است که عمدتا با بهره‌گیری از الگوریتم‌ها و مدل‌های آماری به پردازش، درک و تولید زبان طبیعی می‌پردازد و به عنوان حوزه‌ای به سرعت در حال تکامل، توانسته نحوه تعامل انسان با تکنولوژی را به کلی تغییر دهد. امروزه ماشین‌ها می‌توانند با پردازش پیکره‌ای متنی، زبان انسان‌ها را درک کنند، با آن‌ها ارتباط برقرار کنند و در تصمیم‌گیری‌های روزمره به او کمک کنند. از این نظر، پیشرفت‌های خیره‌کننده این حوزه در جهت بهبود مدل‌های مکالمه‌ای[۱] و چت‌بات‌ها[۲][۳]، ساخت دستیار مجازی کارآمد در امور آموزشی [۴]، مالی [۵] و یا حتی پزشکی [۶]، تولید آثار هنری [۷] و مقالات [۸]، از نمونه دست‌آوردها در این حوزه است. بدیهی است که کاربرد مدل‌های زبانی به موارد ذکر شده محدود نمی‌شود، اما این موارد به خوبی نمایان‌گر تاثیری است که مدل‌های زبانی بر زندگی انسان گذاشته و یا خواهد گذاشت. از مهم‌ترین توانایی‌های یک مدل پردازش زبان که انجام امور زبانی را میسر می‌کند، قابلیت درک معنا[۳] و قصد ارتباطی[۴] گوینده است. این قابلیت، با استفاده از مدل‌سازی معنایی[۵] میسر می‌شود. مدل‌سازی معنایی روشی است برای دست‌یابی به یک بازنمایی از معنا در زبان طبیعی به گونه‌ای که بازنمایی حاصل برای ماشین قابل پردازش باشد و معمولا روابط میان مفاهیم یک زبان را نشان می‌دهد.

---

[1] Conversational Models
[2] Chatbots
[3] Semantic Understanding
[4] Communicative Intent
[5] Semantic Modeling

۱.۱ تعریف موضوع

در ادامه، ابتدا به تعریف بازنمایی معنایی پرداخته و سپس اهمیت و کاربردهای مطالعه ساختاری و مقایسه بازنمایی‌های معنایی را مرور خواهیم کرد.

۱.۱.۱ تعریف بازنمایی معنایی

یک بازنمایی معنایی[۶] صورتی از معنا در زبان طبیعی است که روابط میان مفاهیم مختلف را به گونه‌ای که برای یک مدل کامپیوتری قابل فهم باشد به دست می‌دهد. بازنمایی معنایی معمولا از مدل‌سازی معنایی و یا با بهره‌گیری از دانش انسانی ایجاد می‌شود. بازنمایی معنایی به دست آمده به مدل زبانی کمک می‌کند تا در امور مختلف پردازش زبان‌امور مختلف پردازش زبان[۷] که وابسته به درک معنا هستند موفق عمل کند.

۱.۱.۲ اهمیت موضوع و کاربردها

مطالعه ساختاری و طبقه‌بندی بازنمایی‌های معنایی می‌تواند کاربردهای متعددی داشته باشد که در ادامه به آن می‌پردازیم.

- مطالعه ساختاری یک بازنمایی معنایی به ما این امکان را می‌دهد که به شناخت بهتری نسبت به روش اولیه ساخت آن دست پیدا کنیم. اغلب مطالعات پیشین به بررسی عملکرد روش‌های مدل‌سازی معنا در امور پردازش زبان پرداخته‌اند، این در حالی است که مطالعه ویژگی‌های ساختاری بازنمایی حاصل از یک روش یا مدل ساخت بازنمایی معنایی، می‌تواند اطلاعات مفیدی درباره آن به دست دهد. این موضوع، خصوصا در مورد روش‌های مدل‌سازی توزیعی معنا اهمیت بیشتری پیدا می‌کند؛ چرا که همواره به علت روشن نبودن روند تصمیم‌گیری آن‌ها در امور پردازش زبان مورد انتقاد قرار گرفته‌اند [۹].

- شناخت بهتر ویژگی‌های یک بازنمایی معنایی امکان مقایسه آن با سایر بازنمایی‌ها را به دست می‌دهد. این امر عملا امکان مقایسه روش‌هایی که این بازنمایی‌ها از آن منتج شده‌اند را نیز به دست می‌دهد. در

[6] Semantic Representation
[7] Natural Language Processing Tasks

این صورت، علاوه بر مقایسه روش‌های مختلف مدل‌سازی زبانی با آزمودن قابلیت عملی آن‌ها در امور پردازش زبان، می‌توانیم از خواص بازنمایی‌های آن‌ها برای کشف محدودیت‌های هر یک استفاده کنیم.

- شناخت بهتر بازنمایی‌های معنایی، می‌تواند در تفسیر نتایج حاصل از آزمودن آن‌ها در امور پردازش زبان به ما کمک کند و احتمالا فهم تازه‌ای در جهت بهبود عملکرد یک مدل زبانی به دست دهد.

- درک تفاوت‌های میان بازنمایی‌های معنایی می‌تواند در جهت ترکیب آن‌ها به منظور دست‌یابی به یک بازنمایی با عملکرد بهتر راه‌گشا باشد.

۲.۱ پژوهش‌های پیشین در زمینه مطالعه بازنمایی‌های معنایی

روش‌های متنوعی به منظور ساخت یک بازنمایی معنایی تا کنون ارائه شده‌اند. به طور کلی دو رویکرد در ساخت یک بازنمایی معنایی وجود دارد؛ رویکرد مبتنی بر دانش انسانی[۸] و رویکرد مدل‌سازی توزیعی معنا[۹] [۱۰][۱۱]. در رویکرد اول، وجود و یا عدم رابطه معنایی میان دو مفهوم در زبان بر اساس قضاوت انسانی[۱۰] و یا دانش جمعی[۱۱] تعیین می‌شود که نتیجه آن یک بازنمایی گسسته[۱۲] است. در رویکرد اول، انتخاب روش جمع‌آوری دادگان، روابط معنایی در نظر گرفته شده و منبع جمع‌آوری دادگان مورد استفاده برای ساخت بازنمایی نهایی منجر به ساخت بازنمایی‌های متفاوتی می‌شود. از جمله بازنمایی‌های حاصل از این رویکرد می‌توان کانسپت‌نت[۱۳] و وردنت[۱۴] را نام برد.

در رویکرد دوم، یک مدل آماری آموزش داده می‌شود تا توزیع کلمات در یک پیکره متنی بزرگ را یاد بگیرد و از این طریق یک فضای معنایی پیوسته[۱۵] به دست دهد. در این شرایط، انتخاب روش آماری، ساختار مدل آماری و انتخاب دادگان آموزش بازنمایی معنایی نهایی را دست‌خوش تغییر می‌کند. از جمله مدل‌سازی‌های توزیعی معنایی می‌توان مدل‌های مبتنی بر شمارش کلمات (از جمله ماتریس فراوانی کلمه-معکوس فراوانی

[8] Human-based Approach
[9] Distributional Semantic Modeling
[10] Human Judgment
[11] Crowd Sourcing
[12] Discrete
[13] ConceptNet
[14] WordNet
[15] Continuous

سند[۱۶]، مدل‌های مبتنی بر تعبیه[۱۷] کلمات (مانند Word2Vec [۱۲])، و مدل‌های مبتنی بر بافت[۱۸] کلمات (مانند برت[۱۹] [۱۳] و المو[۲۰] [۱۴]) را نام برد.

روش ساخت یک بازنمایی در ساختار نهایی آن تاثیر داشته [۱۵] و از این رو عملکرد مدل‌های زبانی نیز در امور پردازش زبان با یکدیگر متفاوت است. در حال حاضر، مطالعات بسیاری به منظور سنجش عملکرد مدل‌های زبانی که از این بازنمایی‌ها بهره می‌گیرند ارائه شده است [۱۶]. این در حالی است که اطلاعات اندکی درباره ویژگی‌های ساختاری هر یک از این بازنمایی‌ها داریم [۱۵][۱۷]. هم‌چنین طبقه‌بندی بازنمایی‌های معنایی با استفاده از خواص ساختاری‌شان کمتر مورد توجه قرار گرفته است [۱۵][۱۷][۱۸][۱۹].

در ادامه تعدادی از محدودیت‌های مطالعات پیشین را معرفی می‌کنیم.

- اول از همه، می‌توان به تعداد محدود مطالعات تطبیقی در این زمینه اشاره کرد. برای مثال، به دلیل تفاوت بازنمایی‌های توزیعی و بازنمایی‌های مبتنی بر دانش انسانی از لحاظ پیوستگی فضای آن‌ها، مطالعاتی که این دو دسته را مورد مقایسه قرار دهند بسیار محدودند [۱۰].

- از طرفی دیگر، بازنمایی‌های معنایی ترکیبی در این مطالعات مورد توجه قرار نگرفته‌اند. این در حالی است که بازنمایی‌های ترکیبی در بهبود عملکرد مدل‌های زبانی در امور پردازش زبان تاثیر قابل توجهی دارند. برای مثال، بهره‌گیری از بازنمایی‌های معنایی که علاوه بر داده متنی از داده تصویری نیز بهره می‌گیرند، موجب بهبود عملکرد مدل نهایی در امر تشخیص عبارات مبتنی بر عقل سلیم[۲۱] می‌شود. از طرفی ترکیب بازنمایی‌های معنایی توزیعی و انسانی نیز نتایج خوبی در امر زبانی ذکر شده به دست می‌دهد. با این حال، همان‌طور که گفته شد مطالعه ساختاری این نوع بازنمایی‌ها چندان مورد پژوهش قرار نگرفته است.

- رویکردهای مطالعه بازنمایی‌های معنایی را می‌توان به رویکرد مبتنی بر علم شبکه و رویکرد توزیعی تقسیم کرد. در رویکرد توزیعی، تنوع روش‌ها بسیار محدود است و این موضوع سبب شده که بررسی جنبه‌های مختلف بازنمایی‌های معنایی دشوار شود. از طرفی، رویکردهای مبتنی بر علم شبکه که تا کنون مورد استفاده قرار گرفته‌اند، با وجود در نظر گرفتن سطوح مختلف بازنمایی‌ها، نتوانسته‌اند تفاوت اندازه بازنمایی‌های معنایی را در مقایسه خود لحاظ کنند.

[16] Term Frequency Inverse Document Frequency
[17] Embedding
[18] Context
[19] BERT
[20] ELMo
[21] Commonsense

۳.۱ تعریف مسئله پژوهش و راه‌کار پیشنهادی

در این پژوهش، به انجام مقایسه‌ای جامع میان بازنمایی‌های معنایی متداول در پردازش زبان طبیعی با استفاده از خواص ساختاری آن‌ها می‌پردازیم. به این منظور هفت نوع بازنمایی معنایی در نظر گرفته شده است از جمله آن‌ها می‌توان سه بازنمایی مبتنی بر دانش انسانی و چهار بازنمایی مبتنی بر مدل‌سازی توزیعی معنا را، همان‌طور که در جدول ۱.۱ آمده، نام برد.

جدول ۱.۱: انواع بازنمایی‌های مورد مطالعه در این پژوهش

نام بازنمایی	نوع بازنمایی	توضیح
وردنت	انسانی	روابط هم‌معنایی با استناد به قضاوت انسانی برقرار شده‌اند
فرهنگ‌نامه موبی	انسانی	روابط هم‌معنایی تعریف گسترده‌ای نسبت به وردنت دارند
کانسپت‌نت	انسانی	بیش از سی نوع رابطه معنایی با استناد به قضاوت انسانی و دانش جمعی برقرار شده‌اند
Word2Vec	توزیعی	یک بازنمایی معنایی مبتنی بر مدل‌سازی توزیعی معنا
BERT2Static	توزیعی	اغنای بازنمایی Word2Vec با بهره‌گیری از مدل مبتنی بر بافت برت
VisWord2Vec	توزیعی	اغنای بازنمایی Word2Vec با استفاده از دادگان تصویری
Conceptnet Numberbatch	توزیعی	ترکیب یک بازنمایی انسانی و چند بازنمایی توزیعی از جمله Word2Vec

هدف این است که این مقایسه امکان طبقه‌بندی بازنمایی‌های انسانی و توزیعی را فراهم کند. هم‌چنین، بازنمایی‌های ترکیبی نیز در این مقایسه لحاظ شده‌اند.

به منظور انجام چنین مقایسه‌ای نیاز است که بازنمایی‌های گسسته و پیوسته به یک نوع بازنمایی نگاشت شوند. به این منظور، تمامی بازنمایی‌های معرفی شده در جدول ۱.۱ هر یک به گراف متناظرشان نگاشت می‌شوند. این رویکرد امکانات مختلفی را فراهم می‌کند. اولا با این نگاشت، امکان مقایسه بازنمایی‌های گسسته و پیوسته فراهم می‌شود و به این ترتیب امکان مقایسه بازنمایی‌های انسانی و توزیعی میسر می‌شود. دوما، نگاشت بازنمایی‌ها به گراف، بازنمایی شهودی‌تری نسبت به بازنمایی‌های اولیه اولیه فراهم می‌کند. سوما، در اختیار داشتن گراف هر بازنمایی، به ما امکان استفاده از تعداد زیادی شاخص ساختاری که در علم شبکه معرفی شده‌اند را می‌دهد که هر یک می‌توانند جنبه‌ای از خصوصیات بازنمایی‌های معنایی را آشکار کنند. به علاوه، شاخص‌های شبکه، در سطوح مختلفی از جمله سراسری، میانی و محلی تعریف می‌شوند که این موضوع امکان بررسی بازنمایی‌های معنایی در مقیاس‌های متفاوت را نیز فراهم می‌کند.

در این پژوهش، با بهره‌گیری از شاخص‌های علم شبکه در سطوح سراسری و میانی به مطالعه ساختاری

بازنمایی‌های معنایی می‌پردازیم. به‌علاوه، قصد داریم تا با استفاده از ویژگی‌های ساختاری هر یک از بازنمایی‌های معنایی به طبقه‌بندی آن‌ها بپردازیم. از آن‌جایی که در پژوهش‌های پیشین توجهی به غیرهم‌اندازه بودن بازنمایی‌های معنایی هنگام مقایسه آن‌ها نشده بود، در این پژوهش یک رویکرد آماری را معرفی می‌کنیم که امکان مقایسه گراف بازنمایی‌های معنایی غیرهم‌اندازه را فراهم می‌کند. در این رویکرد، با استفاده از مدل‌های تصادفی متناظر با هر گراف معنایی به طبقه‌بندی آن‌ها می‌پردازیم.

۴.۱ دستاوردهای پژوهش

به طور کلی، دستاوردهای این پژوهش را می‌توان در سه مورد خلاصه کرد:

- در این پژوهش، طبقه‌بندی جامعی از بازنمایی‌های معنایی مورد استفاده در پردازش زبان طبیعی ارائه شده است. از جمله بازنمایی‌هایی که در این پژوهش اولین بار مورد مطالعه و مقایسه با سایر بازنمایی‌ها قرار گرفته‌اند، بازنمایی مبتنی بر داده متنی-تصویری [۲۰]، بازنمایی مربوط به یک مدل غنی شده با گراف دانش [۲۱] و بازنمایی حاصل از یک مدل تعبیه کلمات غنی شده با استفاده از مدل زبانی پویا [۲۲] است. هر یک از موارد ذکر شده کاربردها و عملکردهای مختلفی دارند که مطالعه ساختاری آن‌ها امکان بررسی تاثیر دادگان و روش‌هایی که هر یک به کار گرفته‌اند را در کیفیت بازنمایی نهایی به دست می‌دهد.

- در این پژوهش، رویکردی نوین به منظور مقایسه گراف‌های معنایی غیرهم‌اندازه معرفی شده است. از آنجایی که مقایسه شبکه‌هایی که تعداد رئوس و یال مشابه نداشته باشند دشوار است، با ارائه یک رویکرد آماری سعی کردیم این مسئله را مرتفع کنیم. در این رویکرد، هر شبکه ابتدا با مدل‌های تصادفی خود مقایسه شده و سپس با استفاده از آزمون آماری، شبکه‌های معنایی با یکدیگر مقایسه می‌شوند. از آنجایی که این روش مقایسه غیر مستقیم برای هر نوع گراف دیگری نیز می‌تواند مورد استفاده قرار گیرد، کاربردهای آن به گراف‌های معنایی محدود نمی‌شود.

- پس از اعمال چهارچوب مقایسه پیشنهادی بر گراف‌های معنایی حاصل، دریافتیم که در تمامی بازنمایی‌های انسانی کلماتی که بیشتر در زبان انگلیسی به کار می‌روند، روابط معنایی بیشتری با سایر کلمات دارند. این در حالی است که در بازنمایی‌های معنایی توزیعی، اغلب، کلمات بسیار نادر زبان انگلیسی هستند

که بخش عمده ارتباطات معنایی را به خود اختصاص می‌دهند. به علاوه، دریافتیم که افزودن اطلاعات تصویری به یک بازنمایی توزیعی پایه، موجب کاهش احتمال تشکیل گروه‌های معنایی در آن می‌شود. از طرفی مشاهده کردیم که افزودن اطلاعات موجود در یک بازنمایی انسانی به بازنمایی توزیعی موجب افزایش احتمال تشکیل گروه‌های معنایی می‌شود. همچنین، پس از طبقه‌بندی بازنمایی‌های معنایی مشاهده کردیم که در صورت استفاده از روش ساخت مشابه، لزوما خواص ساختاری بازنمایی‌های حاصل مشابه نخواهند بود. برای مثال، با وجود انسانی بودن روش ساخت بازنمایی‌های وردنت، موبی و کانسپت‌نت، بازنمایی سوم هیچ‌گونه شباهتی به دو بازنمایی دیگر نداشته و به فضاهای معنایی توزیعی شبیه‌تر است. لذا، انتخاب منبع و نوع دادگان می‌تواند تفاوت قابل ملاحظه‌ای میان بازنمایی‌ها ایجاد کند.

۵.۱ ساختار پایان‌نامه

- در فصل دوم، ابتدا به منظور معرفی بازنمایی‌های معنایی مورد مطالعه در این پژوهش، به دسته‌بندی و توضیح انواع بازنمایی‌های معنایی در پردازش زبان طبیعی می‌پردازیم و روش ساخت آن‌ها و کاربرد آن‌ها را شرح می‌دهیم. سپس، به معرفی و مرور رویکردهایی که تا کنون به منظور مطالعه بازنمایی‌های معنایی اتخاذ شده می‌پردازیم، در این قسمت روش‌های استفاده شده در رویکرد غیر شبکه‌ای و هم‌چنین رویکرد مبتنی بر علم شبکه معرفی می‌شوند. سپس، به مروری کلی در زمینه مطالعات انجام شده به منظور مقایسه شبکه‌های غیرهم‌اندازه می‌پردازیم. در انتهای فصل جدول پیشینه تحقیق ارائه می‌شود.

- در فصل سوم، نخست روش نگاشت هر بازنمایی معنایی به گراف متناظر آن شرح داده می‌شود. سپس چهارچوب پیشنهادی جهت مطالعه و مقایسه شبکه‌های معنایی حاصل، معرفی خواهد شد. این چهارچوب از دو بخش مقایسه در مقیاس سراسری و در مقیاس میانی تشکیل می‌شود که هر بخش شامل شاخص‌های مربوطه است. تعریف و روش محاسبه هر شاخص در این فصل آورده شده است. سپس رویکرد نوین مقایسه شبکه‌های غیرهم‌اندازه معنایی معرفی می‌شود. این رویکرد، مقایسه‌ای است غیرمستقیم، که از آزمون آماری بهره می‌گیرد.

- فصل چهارم، دربرگیرنده نتایج حاصل از به کارگیری چهارچوب معرفی شده در فصل قبلی است. ابتدا شبکه‌های معنایی حاصل از هر نگاشت ارائه می‌شوند. سپس، هر یک از شاخص‌های سراسری بر گراف‌های

معنایی اعمال شده و نتایج گزارش می‌شوند. پس از آن، با استفاده از نتایج این شاخص‌ها و انجام مقایسه آماری گراف‌های معنایی به طبقه‌بندی این بازنمایی‌ها می‌پردازیم. در نهایت، فصل چهارم را با ارائه نتایج معیارهای سطح میانی به پایان می‌رسانیم.

- در فصل پنجم و آخر، ابتدا به بحث درباره نتایج گزارش شده در فصل چهارم پرداخته می‌شود. در این قسمت ابتدا به جمع‌بندی دست‌آوردهای رویکرد نوین آماری جهت مقایسه شبکه‌های غیرهم‌اندازه می‌پردازیم. در مرحله بعدی، شباهت‌ها و تفاوت‌های شبکه‌های مورد مطالعه مورد بررسی قرار می‌گیرد. همچنین، ارتباط این نتایج با روش ساخت بازنمایی مربوطه و مورد بحث قرار خواهد گرفت. در نهایت، محدودیت‌های چهارچوب مورد استفاده در این مطالعه و دیگر محدودیت‌های عملی که با آن مواجه بودیم معرفی شده و تحقیق جاری را با ارائه پیشنهاداتی جهت مطالعه بهتر بازنمایی‌های معنایی و پژوهش‌های بیشتر به پایان می‌بریم.

فصل ۲

مروری بر مطالعات انجام‌شده

۱.۲ مقدمه

در این فصل مروری بر پژوهش‌های انجام شده به منظور مطالعه و مقایسه بازنمایی‌های معنایی ارائه می‌شود. به این منظور، ابتدا به معرفی و دسته‌بندی انواع مدل‌سازی‌های رایج معنایی می‌پردازیم و به طور کلی روش‌ها و دادگان مورد استفاده در هر مدل‌سازی را مرور خواهیم کرد. سپس، رویکردهایی که تا کنون در مبحث مطالعه خواص بازنمایی‌های معنایی اتخاذ شده است برای خواننده شرح داده می‌شود. در این قسمت با توجه به پژوهش‌های قبلی، دو رویکرد کلی را می‌توان برشمرد؛ رویکرد نخست مبتنی بر استفاده از ابزارهای غیر شبکه‌ای و رویکرد دوم مبتنی بر ابزارهای شبکه‌ای است.

پس از مرور روش‌های مطالعه بازنمایی‌های معنایی، با توجه به غیرهم‌اندازه بودن بازنمایی‌های معنایی مورد مطالعه و در نتیجه غیر هم‌اندازه بودن شبکه‌های معنایی حاصل از این بازنمایی‌ها، به مروری بر روش‌های مقایسه شبکه‌های غیرهم‌اندازه می‌پردازیم.

با توجه به محتوای این فصل، در انتها جدول پیشینه تحقیق نیز آورده شده است که حاکی از محدودیت مطالعات پیشین در جامعیت از لحاظ در نظر گرفتن انواع بازنمایی‌های معنایی است. به علاوه این جدول نشان می‌دهد که رویکرد آماری ارائه شده در تحقیق جاری، برای اولین بار معرفی شده و چهارچوب‌های مقایسه شبکه‌های غیرهم‌اندازه که قبلا معرفی شده‌اند، از این نوع مقایسه، بهره نبرده‌اند.

۱.۱.۲ تعاریف و مبانی نظری

معنا با توجه به نبود تعریفی مشخص برای معنا که بر آن توافق همگانی وجود داشته باشد، تعریف آن آسان نیست. در این پژوهش، تعریف ارائه شده توسط بندر و کولر [1] برای معنا را که به شرح زیر است در نظر می‌گیریم؛

> «معنا را می‌توان به صورت رابطه $M \subseteq E \times I$ تعریف کرد که شامل جفت‌های (e, i) است که در آن e عباراتی از زبان طبیعی و i یک قصد ارتباطی است که e می‌تواند فراخواننده آن باشد».[۲۳]

با این تعریف، معنای یک عبارت از زبان طبیعی درباره چیزی در جهان واقعی و یا جهان ذهن است و لزوما به روابط میان مفاهیم در زبان طبیعی محدود نمی‌شود.

بازنمایی معنایی یک بازنمایی معنایی صورتی از معنا در زبان طبیعی است به صورتی که برای یک مدل کامپیوتری قابل فهم باشد.

بافت بافت یا بافتار نیز ممکن است تعاریف مختلفی داشته باشد، هر چند در این تحقیق، بافت برای هر کلمه، کلماتی هستند که به همراه کلمه مورد نظر در متن ظاهر شده باشند.

بازنمایی ایستا بازنمایی ایستای [2] معنایی، نوعی بازنمایی است که در آن معانی مختلف یک کلمه در بافتارهای متفاوت لحاظ نشده و هر کلمه به یک بازنمایی ثابت نگاشت شده است. به بیان دیگر، در این نوع بازنمایی برای کلمات تنها معنای ثابت [3] (مستقل از بافتار [4]) لحاظ شده است [۱۱].

بازنمایی پویا برخلاف بازنمایی ایستا، یک بازنمایی پویا یا غیر یک‌ریخت [5]، علاوه بر معنای ثابت کلمات، معنای موقعیتی [6] یا وابسته به بافت [7] آن‌ها را نیز در نظر می‌گیرد. واضح است که بازنمایی نهایی ثابت نیست و با

[1] Bender and Koller
[2] Static
[3] Standing Meaning
[4] Context-independent
[5] Nonisomorphic
[6] Occasion Meaning
[7] Context-dependent

توجه به متن می‌تواند تغییر کند [۱۱].

۲.۲ معرفی و دسته‌بندی انواع بازنمایی‌های معنایی مورد استفاده در پردازش زبان طبیعی

بازنمایی‌های معنایی را، از نظر روش ساخت، می‌توان به دو دسته کلی تقسیم کرد: بازنمایی‌های مبتنی بر دانش انسانی و بازنمایی‌های مبتنی بر یادگیری ماشین (توزیعی). در ادامه به تعریف این دو نوع بازنمایی معنایی می‌پردازیم و با جزییات بیشتری به روش ساخت هر یک از آن‌ها خواهیم پرداخت. لازم به ذکر است که این دو دسته هر کدام انواع مختلفی از روش‌های ساخت یک بازنمایی معنایی را در بر می‌گیرند که در ادامه توضیح داده خواهد شد.

۱.۲.۲ بازنمایی‌های معنایی مبتنی بر دانش انسانی

همان‌گونه که از نام این رویکرد برمی‌آید، این دسته از بازنمایی‌های معنایی عموما با بهره‌گیری از دانش انسانی ساخته می‌شوند.

به این منظور، روابط معنایی که ممکن است میان مفاهیم یک زبان وجود داشته باشد، از قبل تعریف شده و بر این اساس، یک داور انسانی درباره عدم و یا وجود نوعی ارتباط میان دو مفهوم تصمیم می‌گیرد. گاهی گروه داوران انسانی عموما متشکل از افراد متخصص زبان[8] هستند. برای مثال، در دو بازنمایی معنایی وردنت [9] [۲۴] و فرهنگ معنایی موبی [10] [۲۵] رابطه در نظر گرفته شده میان مفاهیم رابطه هم‌معنایی است و عموما افراد متخصص زبان در ساخت بازنمایی نهایی مشارکت دارند.

بازنمایی وردنت به صورت مجموعه‌ای از زیرمجموعه‌های معنایی ارائه شده است. هر زیرمجموعه، شامل کلماتی است که بر اساس قضاوت انسانی، با یکدیگر رابطه معنایی[11] دارند. این بازنمایی در بهبود مدل‌های

[8] Lexicographers /Linguists
[9] WordNet
[10] Moby
[11] Semantic Relation

زبانی[۱۲] برای امور تحلیل احساسات متن[۱۳] [۲۶]، ابهام‌زدایی معنایی کلمات[۱۴] [۲۷] و تشابه‌یابی متون [۲۸] به کار می‌رود. گاهی در برخی بازنمایی‌ها مانند کانسپت‌نت [۱۵] [۲۹]، ممکن است علاوه بر دانش افراد متخصص از دانش جمعی گویشوران زبان[۱۶] بهره گرفته شود و یا حتی قسمتی از مرحله جمع‌آوری داده به صورت خودکار[۱۷]، مثلا با بازی‌های هدف‌دار[۱۸]، انجام شود. بازنمایی نهایی کانسپت‌نت به صورت مجموعه‌ای از جفت مفاهیم[۱۹] ارائه شده است. هر جفت بیانگر وجود یک رابطه معنایی میان دو مفهوم متناظر است. انواع روابط معنایی موجود در این بازنمایی بسیار متنوع‌تر از دو بازنمایی قبلی است و به همین علت، استفاده از کانسپت‌نت برای بهبود عملکرد مدل‌های زبانی در اموری که نیاز به دسترسی به فهم متعارف[۲۰] انسانی دارند نتایج مطلوبی به دست می‌دهند. از جمله این امور می‌توان امر تمییز جملات مبتنی بر فهم متعارف از جملات بی‌معنا[۲۱] [۳۰] را نام برد.

قابل ذکر است که بازنمایی‌های معنایی مبتنی بر دانش انسانی، ممکن است بیش از یک زبان را در بر بگیرند، مانند کانسپت‌نت و بابل‌نت[۲۲] [۳۱].

۲.۲.۲ بازنمایی‌های معنایی مبتنی بر مدل‌سازی توزیعی معنا

رویکرد دیگر دست‌یابی به یک بازنمایی معنایی، استفاده از روش‌های توزیعی مدل‌سازی معنا[۲۳] است. در مدل‌سازی توزیعی معنا، هدف، یادگیری معنای عبارات زبانی با استفاده از یک پیکره متنی[۲۴] است [۱۱]. در این رویکرد، بافت[۲۵] کلمات در یک پیکره، می‌تواند در فهم معنای کلمات راه‌گشا باشد. به بیان دیگر، طبق فرضیه‌ای موسوم به فرضیه توزیعی[۲۶] [۳۲] هرچه دو عبارت زبانی بیشتر در یک بافت مشترک ظاهر شوند، احتمال شباهت معنایی آن‌ها نیز بیشتر است. تکامل روش‌های مدل‌سازی توزیعی معنا در طول زمان را می‌توان به شکل زیر تشریح

[12] Language Models
[13] Sentiment Analysis
[14] Word Sense Disambiguation
[15] ConceptNet
[16] Crowd Sourcing
[17] Automatic
[18] Games with Purpose
[19] Assertion
[20] Commonsense
[21] Commonsense Reasoning
[22] BabelNet
[23] Distributional Semantic Modelling
[24] Textual Training Corpus
[25] Context
[26] Distributional Hypothesis

کرد.

۱.۲.۲.۲ مدل‌های مبتنی بر شمارش کلمات

این مدل‌های بر اساس شمارش دفعات باهم‌آیی کلمات در متون موجود در پیکره آموزش ساخته می‌شوند. برخی از این‌ها مانند ماتریس فراوانی کلمه-معکوس فراوانی سند[۲۷] مدل شمارشی ساده هستند و برخی مانند روش تحلیل معنایی پنهان[۲۸] [۳۳]، از تجزیه یک ماتریس شمارشی به دست می‌آیند.

۲.۲.۲.۲ مدل‌های مبتنی بر تعبیه کلمات

با معرفی مدل Word2Vec [۱۲] در سال ۲۰۱۳ مدل‌های شمارشی جای خود را به مدل‌های تعبیه کلمات دادند. در این روش، یک مدل شبکه عصبی[۲۹] وظیفه یادگیری یک بازنمایی معنایی از کلمات موجود در پیکره آموزش را بر عهده دارد. به این منظور، ابتدا یک پنجره با طول مشخص از قبل تعیین می‌شود و در هر بار حرکت پنجره روی متن آموزش، یک کلمه به عنوان هدف انتخاب شده و دیگر کلمات پنجره به عنوان بافت کلمه هدف به عنوان ورودی به شبکه عصبی داده می‌شود. شبکه عصبی مذکور باید بتواند کلمه هدف را پیش‌بینی کند[۳۰]. در این روش، وزن‌های لایه پنهان[۳۱] شبکه عصبی، بازنمایی برداری کلمه پیش‌بینی شده را به دست می‌دهند. بازنمایی معنایی نهایی یک فضای n بعدی است که در آن n تعداد نورون‌های[۳۲] لایه پنهان است. پس از معرفی Word2Vec روش‌های تعبیه کلمات بسیاری، از جمله GloVe [۳۴] و FastText [۳۵] ارائه شدند.

۳.۲.۲.۲ مدل‌های مبتنی بر بافت

بر خلاف مدل‌های پیشین که هر کلمه را به یک بردار ثابت در فضای معنایی نهایی نگاشت می‌کنند، مدل‌های این دسته از بازنمایی توزیعی معنا، نگاشت یک به یک از کلمات به بردار معنایی متناظرشان ندارند. به بیان دیگر،

[27] TF-IDF
[28] Latent Semantic Analysis
[29] Neural Netwok Model
[۳۰] توضیح داده شده مربوط به معماری کیسه لغات پیوسته است. Word2Vec معماری دیگری به نام Skip-gram نیز دارد که در اینجا آورده نشده است.
[31] Hidden Layer
[32] Neuron

بازنمایی برداری یک کلمه، تابعی از کلماتی است که با آن کلمه در جمله آمده‌اند. لذا با تغییر کلمات یک جمله، بردار کلمه نیز تغییر می‌کند. با این توضیح، مشخص می‌شود که بازنمایی نهایی حاصل از این نوع مدل‌سازی توزیعی معنا یک‌شکل[۳۳] و یا ایستا[۳۴] نبوده و درواقع یک بازنمایی پویا[۳۵] است. به همین علت، این نوع بازنمایی قادر به در نظر گرفتن معانی متفاوتی[۳۶] است که یک کلمه در بافتار مختلف ممکن است داشته باشد. مدل المو [۱۴] که از معماری حافظه طولانی کوتاه‌مدت[۳۷] بهره می‌برد و دو مدل برت [۱۳] و جی‌پی‌تی[۳۸] [۳۶] که بر پایه معماری مبدل[۳۹] ساخته شده‌اند، از نمونه مدل‌های زبانی هستند که بازنمایی معنایی مبتنی بر بافت به دست می‌دهند.

واضح است که در تمامی روش‌های مدل‌سازی توزیعی معنا که در این قسمت بررسی شد، یادگیری بازنمایی معنایی نهایی، بدون ناظر[۴۰] صورت می‌گیرد. به علاوه، با توجه به توضیحات ارائه شده، می‌توان گفت اینکه یادگیری روابط معنایی از پیکره آموزش چگونه انجام می‌شود، مستقیما به الگوریتم یادگیری و یا ساختار شبکه عصبی مربوطه وابسته است.

۳.۲.۲ بازنمایی‌های معنایی ترکیبی

در این بخش، به مرور کلی بازنمایی‌های ترکیبی می‌پردازیم. منظور از ترکیبی در این بخش، ترکیب دادگان آموزش متفاوت و یا روش‌های ساخت متفاوت برای بهبود یا اختصاصی‌سازی یک بازنمایی معنایی است.

۱.۳.۲.۲ بازنمایی ترکیبی آموزش دیده بر متن و دادگان ادراک حسی

اولین روش ساخت یک بازنمایی ترکیبی، بهره‌گیری از دادگان مربوط به ادراک حسی[۴۱] است. تمامی بازنمایی‌های معنایی که تا اکنون معرفی شدند، برای تعریف معنای یک کلمه مشخص، تکیه بر روابط موجود میان آن کلمه و سایر کلمات می‌کردند. این در حالی است که اگر معنای کلمات تنها بر اساس سایر کلمات تعریف شوند، به

[33] Isomorphic
[34] Static
[35] Dynamic
[36] Polysemy
[37] Long Short-term Memory
[38] GPT
[39] Transformer
[40] Unsupervised
[41] Sensory Perception

تعاریفی دوار[۴۲] می‌رسیم [۳۷]. به همین علت، برقراری ارتباط میان مفاهیم یک زبان و موجودیت‌ها در جهان واقعی می‌تواند راه حلی برای گریز از این نوع ارجاع دوار باشد [۱۱]. به فرآیند برقراری ارتباط میان یک بازنمایی معنایی و عناصر جهان واقعی، زمینی کردن[۴۳] گفته می‌شود. همان‌طور که پیش‌تر درباره تعریف معنا، ارائه شده توسط بندر و کولر [۲۳] پرداختیم، معنای مفاهیم چیزی فراتر از صرف روابط میان کلمات در یک زبان است. آن‌ها ادعا می‌کنند که مدل‌های معنایی هم‌چون برت و جی‌پی‌تی، علی‌رغم توان آماری بسیار بالا در پردازش زبان و عملکرد قابل توجه در امور مختلف زبانی، به دلیل اینکه تنها با استفاده از داده متنی آموزش داده شده‌اند، دسترسی به تمام آنچه که معنا می‌خوانیم ندارند. آن‌ها ادعا می‌کنند که داده متنی به‌تنهایی، لزوما تمامی روابط معنایی که در جهان واقعی وجود دارد را در برنمی‌گیرد. بندر و کولر راه حلی این مسئله را زمینی کردن یادگیری معنا در این مدل‌ها با بهره‌گیری از دادگان مبتنی بر ادراک تصویری می‌دانند. امرسون [۱۱] انواع دیگری از حواسی که تا به امروز از آنها برای مدل‌سازی معنایی استفاده شده از جمله بویایی [۳۸] و شنیداری [۳۹] را نیز نام می‌برد.

با توجه به سختی لحاظ کردن این نوع از دادگان در بازنمایی معنایی، بازنمایی ترکیبی متن-تصویر تا به امروز از بقیه انواع رایج‌تر بوده است. این نوع بازنمایی، علاوه بر بهبود عملکرد مدل‌های زبانی در امر بازشناسی جملات مبتنی بر فهم متعارف [۲۰]، در امور مربوط به پردازش تصویر[۴۴] مانند سوال و جواب دیداری[۴۵] [۴۰] و یا توضیح نویسی تصویر[۴۶] [۴۱] نیز کاربرد بسیاری دارد و به همین دلیل پیاده‌سازی‌های متعددی نیز برای آموزش این نوع بازنمایی موجود است.

۲.۳.۲.۲ بازنمایی ترکیبی حاصل از مدل معنایی توزیعی و گراف دانش

دومین روش رایج ساخت یک بازنمایی معنایی ترکیبی، استفاده از یک پایگاه داده ساختارمند به منظور بهبود فضای معنایی حاصل از یک مدل معنایی توزیعی است. از آن‌جایی که یادگیری بازنمایی معنایی معمولا از یک پیکره متنی بدون ساختار اتفاق می‌افتد، استفاده از یک مجموعه داده ساختارمند مانند یک گراف دانش، ممکن است روابط معنایی موجود در بازنمایی اولیه و در نتیجه عمل‌کرد آن را بهبود ببخشد. از گراف‌های دانشی که به این منظور مورد استفاده قرار می‌گیرند می‌توان کانسپت‌نت [۲۱] و ویکیپدیا[۴۷] [۴۲] را نام برد.

[42] Circular Definitions
[43] Language Grounding
[44] Image Processing
[45] Visual Question Answering
[46] Image Captioning
[47] Wikipedia

۳.۳.۲.۲ بازنمایی ترکیبی حاصل از مدل معنایی توزیعی ایستا و پویا

اخیرا به منظور بهره‌گیری از مزایای هر دو نوع بازنمایی ایستا (به صرفه بودن از لحاظ محاسباتی) و بازنمایی پویا (در نظر گرفتن بافتار)، مدل‌های ترکیبی از این بازنمایی‌ها ارائه شده است [۴۳] [۲۲]. نشان داده شده که بازنمایی‌های نهایی در امر محاسبه شباهت کلمات از مدل‌های ایستای پایه مانند Word2Vec و FastText بهتر عمل می‌کنند [۲۲][۴۴].

۳.۲ تاریخچه مطالعه بازنمایی‌های معنایی

در این بخش به مرور پژوهش‌های پیشین در رابطه با مطالعه بازنمایی‌های معنایی و مقایسه آن‌ها می‌پردازیم و دو دسته کلی از رویکردها را در نظر می‌گیریم. دسته نخست، مطالعاتی هستند که از هر نوع ابزاری به غیر از ابزار مبتنی بر علم شبکه برای مطالعه بازنمایی‌های معنایی استفاده کرده‌اند. با توجه به گوناگونی روش‌ها و نبود چهارچوب‌های متداول، در این تحقیق، دسته نخست را رویکردهای مبتنی بر ابزار غیر شبکه‌ای می‌نامیم. دسته دوم نیز مطالعاتی را در بر می‌گیرد که از هرگونه ابزار مبتنی بر علم شبکه برای مطالعه و یا مقایسه بازنمایی‌های معنایی بهره برده‌اند.

قابل ذکر است که برخی از این مطالعات مقایسه‌ای[۴۸] هستند و برخی صرفا به مطالعه خصوصیات یک بازنمایی مشخص پرداخته‌اند.

۱.۳.۲ رویکردهای غیر از شاخص‌های علم شبکه

همان‌طور که گفته شد، گوناگونی انواع معیارها در این دسته زیاد است. در این دسته از مطالعات، بازنمایی‌های انسانی و بازنمایی‌های توزیعی معمولا جدا از هم مطالعه شده و مقایسه جامعی میان این دو صورت نگرفته است.

[48] Comparative

۱.۱.۳.۲ معیار ناهم‌گونی و پراکندگی برداری (توزیعی)

یکی از روش‌های رایج مطالعه بازنمایی‌های معنایی توزیعی، مطالعه میزان ناهمگونی[۴۹] در بازنمایی آن‌هاست که در واقع نشان‌دهنده میزان گوناگونی یک بازنمایی در جهت‌های مختلف است. به بیان دیگر، ناهمگونی معیاری است از اینکه که توزیع بازنمایی کلمات در همه جهات یک بازنمایی معنایی تا چه اندازه غیر نرمال انجام شده است.

برای محاسبه میزان ناهمگونی، عمدتا از متوسط شباهت کسینوسی میان تعدادی بردار تصادفی و یا روشی مبتنی بر تحلیل مولفه‌های اصلی[۵۰] ماتریس بازنمایی کلمات استفاده می‌شود [۱۹].

در سال ۲۰۱۷ میمنو و تامپسون[۵۱] نشان می‌دهند که در مدل تعبیه کلمات اسکیپ‌گرام[۵۲] بازنمایی نهایی، بر خلاف انتظار، ناهمگون است. تصور پیشین این بود که در بازنمایی برداری کلمات، احتمالا مفاهیم با توجه به شباهت معنایی‌شان حوزه‌های معنایی[۵۳] متفاوتی در سراسر فضای معنایی تشکیل می‌دهند اما مشاهدات این پژوهشگران نشان داد که در بازنمایی حاصل از اسکیپ‌گرام، غالب بردار کلمات در جهت یک بردار میانگین سازمان یافته‌اند. این یافته تا پیش از این در بازنمایی فضای برداری با استفاده از روش کاهش بعد تعبیه تصادفی همسایگان با توزیع t[۵۴] نمایان نشده بود [۱۷].

اتایاراج[۵۵] در [۴۴] نشان می‌دهد که بازنمایی‌های مبتنی بر بافت (از جمله المو و برت) بسیار غیرهمگون هستند؛ لذا بردار کلمات در فضای معنایی حاصل بخش مخروطی کوچکی را اشغال می‌کنند. به طور کلی این ناهمگونی در لایه‌های آخر بیشتر می‌شود که نشان می‌دهد حساسیت به بافت کلمات[۵۶] در لایه آخر مدل‌های مبتنی بر بافت بیشتر است.

در سال ۲۰۲۲، پیله‌ور و همکاران مطالعه ناهمگونی بازنمایی‌های معنایی را به منظور مقایسه مدل‌های زبانی تک زبانه و چند زبانه انجام می‌دهند [۱۹]. نتایج پژوهش آن‌ها نشان می‌دهد که فضاهای چندزبانه مبتنی بر بافت مشابه فضاهای معنایی تک‌زبانه ناهم‌گونی بسیار بالایی دارند اما این ناهم‌گونی‌ها در لایه‌های مختلف تفاوت قابل توجهی ندارند. آن‌ها همچنین نشان می‌دهند که ناهم‌گونی در بازنمایی تمامی زبان‌ها قابل مشاهده است.

[49] Anisotropy
[50] Principle Component Analysis
[51] Mimno and Thompson
[52] Skip-gram
[53] Semantic Field
[54] t-Distributed Stochastic Neighbor Embedding (t-SNE)
[55] Kawin Ethayarajh
[56] Context-sensitivity

از طرفی، پیله‌ور و همکاران مشاهده کرده‌اند که افزایش همگونی فضای معنایی چندزبانه، عملکرد آن را در امر بازنمایی شباهت معنایی[۵۷] بهبود می‌بخشد.

مطالعه مشابهی توسط بدر و همکاران [۴۵] با استفاده از مفهوم ناهمگونی صورت گرفته است. آن‌ها به منظور بازشناسی خصوصیات تعبیه آکوستیک کلمات، به بررسی نحوه توزیع بردار اصوات کلمات و به بیان دیگر، به بررسی خواص بازنمایی اصوات می‌پردازند. با بهره‌گیری از مفهوم ناهمگونی، آن‌ها چگونگی و نحوه تاثیر ساختمان مدل آماری انتخاب شده بر کیفیت بازنمایی نهایی مطالعه می‌کنند.

در سال ۲۰۱۸، چاندراهاس[۵۸] و همکاران، به منظور مطالعه هندسی مدل‌های تعبیه گراف دانش[۵۹]، دو معیار تمرکز و پراکندگی بردارها را معرفی کرده و از آن‌ها به منظور مقایسه روش‌های تعبیه گراف دانش استفاده می‌کنند [۱۵]. آن‌ها هم‌چنین تاثیر تغییر ویژگی‌های هندسی این فضاها در عملکرد نهایی آن‌ها را مورد بررسی قرار می‌دهند. معیار تمرکز (مخروطی بودن[۶۰]) شباهت بسیار زیادی به معیار ناهم‌گونی که پیش‌تر ارائه شد دارد. این معیار نیز بر حسب شباهت کسینوسی تعریف می‌شود و میزان تجمع بردارهای بازنمایی نهایی را نشان می‌دهد. در مقابل، میزان پراکندگی بردارهای بازنمایی با استفاده از معیار پراکندگی که آن هم با استفاده از شباهت کسینوسی تعریف می‌شود محاسبه می‌شود. چاندراهاس و همکاران مشاهده کردند که روش تعبیه دانش افزایشی[۶۱] منجر به یک بازنمایی با تمرکز برداری پایین و پراکندگی بالا می‌شود. این در حالی است که بازنمایی حاصل از روش ضربی[۶۲]، تمرکز برداری بالا و پراکندگی کمی دارد.

۲.۱.۳.۲ کاهش بعد و تصویرسازی

شن و همکاران[۶۳] در سال ۲۰۲۰، با هدف درک بهتر ساختار فهم متعارف، به بررسی ساختار گراف دانش کانسپت‌نت می‌پردازند [۴۶] و تمرکز آن‌ها در این پژوهش انواع و توزیع روابط معنایی از پیش‌تعریف شده در کانسپت‌نت است. به این منظور، آن‌ها با استفاده از روش‌هایی مانند محاسبه هم‌پوشانی کلمات در انواع روابط به محاسبه شباهت‌های موجود میان این روابط می‌پردازند. هم‌چنین، با تعبیه کانسپت‌نت در فضای برداری و

[57] Semantic Similarity
[58] Chandrahas
[59] Knowledge Embedding Models
[60] Conicity
[61] Additive
[62] Multiplicative
[63] Ke Shen

تصویرسازی بردارهای حاصل با روش کاهش بعد تعبیه تصادفی همسایگان با توزیع t، به شناسایی خوشه‌های معنایی در کانسپت‌نت می‌پردازند.

تصویرسازی یک بازنمایی با کاهش بعد به منظور مطالعه خصوصیت‌های آن برای مدل‌های معنایی ایستا امکان پذیر است اما انجام این کار برای مدل‌های پویا مانند برت نتایج قابل قبولی به دست نمی‌دهد. به همین علت، دوسی[۶۴] و همکاران در سال ۲۰۲۲، به منظور شناسایی سوگیری‌های جنسیتی در بازنمایی معنایی حاصل از برت یک روش کاهش بعد معرفی می‌کنند که عملکرد بهتری از روش‌های کاهش بعد پایه دارد. برای شناسایی سوگیری‌های جنسیتی، آن‌ها از رویکردی ترکیبی مبنی بر تحلیل مولفه‌های اصلی و یک دسته‌بند خطی بردار پشتیبان[۶۵] بهره می‌گیرند [۴۷].

چانگ[۶۶] و همکاران [۱۸] به منظور مطالعه فضاهای معنایی چندزبانه روند تقریبا مشابهی را از لحاظ استفاده از تصویرسازی کاهش بعد داده شده پیش می‌گیرند. آن‌ها ابتدا نشان می‌دهند که چگونه در یک بازنمایی واحد فضای معنایی مستقل هر زبان و فضای معنایی مشترک زبانی لحاظ می‌شود. برای این کار با استفاده از تجزیه مقادیر منفرد[۶۷]، یک تبدیل همگر[۶۸] حول میانگین بازنمایی‌های مبتنی بر بافت کلمات هر زبان انجام می‌دهند و بازنمایی اولیه را به هشتاد و هشت زیرفضا[۶۹] تبدیل می‌کنند. سپس با استفاده از تخصیص پنهان دیریکله[۷۰] به شناسایی محورهای حساس به زبان[۷۱] (از جمله خانواده زبان‌ها[۷۲]) و محورهای بی‌تفاوت به زبان[۷۳] (از جمله مقوله‌های نحوی[۷۴]) می‌پردازند. با توجه به این‌که کاهش بعد در این پژوهش به صورت معمول انجام نمی‌شود، تصویرسازی نهایی اطلاعات بیشتری به منظور مطالعه بازنمایی چندزبانه به دست می‌دهد.

[64] Michele Dusi
[65] Linear Support Vector Classifier
[66] Chang
[67] Singular Value Decomposition
[68] Affine Transformation
[69] Subspace
[70] Latent Dirichlet Allocation
[71] Language Sensitive Axes
[72] Language Families
[73] Language-neutral Axes
[74] Part of Speech

۲.۳.۲ رویکردهای مبتنی بر علم شبکه

استفاده از ابزارهای مبتنی بر علم شبکه بیشتر به منظور ساخت مدل‌های پردازش زبان که قادر به انجام امور مختلف زبانی هستند، صورت می‌گیرد. از جمله این امور می‌توان دسته‌بندی متون [۴۸][۴۹] و ابهام‌زدایی معنایی کلمات [۵۰] را نام برد. به بیان دیگر، به رغم امکانات متعدد رویکرد مبتنی بر شبکه، مطالعاتی که تا به امروز از این ابزارها به منظور مطالعه و یا مقایسه بازنمایی‌های معنایی بهره برده باشند انگشت شمار هستند.

عمده تحقیقات صورت گرفته قبل از سال ۲۰۱۵ در مطالعه بازنمایی‌های معنایی با استفاده از شاخص‌های مبتنی بر علم شبکه، پژوهش‌هایی غیر تطبیقی هستند و تمرکز آن‌ها درک خواص زبان و چگونگی یادگیری آن توسط انسان‌ها، با استفاده از شاخص‌های شبکه است.

کانچو[۷۵] و همکاران در سال ۲۰۰۱، با ساخت یک گراف مبنی بر باهم‌آیی کلمات در متن، ویژگی‌های آن را متفاوت از گراف‌های تصادفی[۷۶] و آن را یک شبکه پیچیده دانسته‌اند. آن‌ها نشان داده‌اند که این گراف‌ها متوسط فاصله کوچک (۲ الی ۳) دارند و نحوه توزیع درجات در آن‌ها مستقل از قیاس[۷۷] است [۵۱]. در سال ۲۰۰۹، چودهوری[۷۸] و همکاران، به منظور مطالعه پیچیدگی‌های زبانی آن را به مثابه یک سازوکار فیزیکی[۷۹] در نظر می‌گیرند که پویا و تطبیق‌پذیر[۸۰] است. آن‌ها برای این سازوکار سه سطح محلی[۸۱]، میانی[۸۲] و سراسری[۸۳] را در نظر می‌گیرند و انواع رویکردهایی که منجر به مدل‌سازی زبان به شکل یک شبکه پیچیده می‌شود را بررسی می‌کنند. هم‌زمان فوکس[۸۴] و همکاران [۵۲]، به بررسی تغییرات شبکه زبانی حین یادگیری زبان می‌پردازند. آن‌ها نشان میدهند که با افزایش اندازه شبکه زبانی، ضریب خوشه‌بندی ابتدا کاهش پیدا کرده و زمانی که تعداد رئوس به هزار تا ده‌هزار می‌رسد این شاخص به یک مقدار حداقلی می‌رسد اما بعد از آن شروع به افزایش می‌کند. به طور مشابه، سوله[۸۵] و همکاران نیز با استفاده از علم شبکه، به مدل‌سازی و بررسی روند تکامل زبان[۸۶] می‌پردازند [۵۳].

[75] Cancho
[76] Random Networks
[77] Scale-free Distribution
[78] Choudhury
[79] Physical System
[80] Dynamic Adaptive Complex System
[81] microscopic
[82] Mesoscopic
[83] Macroscopic
[84] Fuks
[85] Sole
[86] Language Evolution

در سال ۲۰۱۵، بیمن[87] و همکاران، به منظور مقایسه بازنمایی حاصل از مدل‌های زبانی n-گرام و زبان طبیعی، از دو معیار ضریب خوشه‌بندی و تحلیل موتیف[88] بهره می‌گیرند. آن‌ها مشاهده کردند که این دو معیار تفاوت قابل توجهی میان اسناد تولید شده توسط n-گرام‌ها و اسناد تولید شده توسط انسان نشان می‌دهند [۵۴].

در سال ۲۰۱۹ ورمیف[89] و همکاران چهارچوبی به منظور مقایسه بازنمایی‌های مبتنی بر دانش انسانی و بازنمایی‌های حاصل از مدل‌سازی توزیعی (یادگیری ماشین) ارائه کردند [۱۰]. در چهارچوب پیشنهادی، تمامی بازنمایی‌ها به شبکه‌های معنایی تبدیل می‌شدند و از این طریق مقایسه انواع بازنمایی‌های معنایی امکان‌پذیر می‌شد. نگاشت یک بازنمایی مبتنی بر دانش انسانی به یک گراف امری ساده است اما نگاشت یک بازنمایی برداری به یک گراف به سادگی امکان‌پذیر نبود. نویسندگان به منظور انجام این نگاشت، میزان شباهت کسینوسی میان جفت بردارهای کلمات را اندازه گرفته و سپس، با استفاده از یک آستانه تشابه از پیش تعریف شده، به جفت کلماتی که بیشتر از آستانه به هم شبیه بودند یال اختصاص داده‌اند. به منظور مقایسه، بازنمایی‌های حاصل از وردنت، فرهنگ موبی و Word2Vec با استفاده از شاخص‌های سراسری مختلفی مطالعه شدند.

در سال ۲۰۲۰، ینیچلیک و همکاران[۵۵]، برای مطالعه چگونگی لحاظ پدیده چندمعنایی در مدل زبانی برت، از تبدیل این بازنمایی به یک گراف استفاده کردند. به این منظور، آن‌ها یک نمونه‌ای از کلمات را انتخاب کرده و برای هر جفت کلمه در این نمونه، شباهت کسینوسی بردارهای متناظر آن‌ها را محاسبه می‌کنند. با استفاده از این مقادیر و یک آستانه تشابه از پیش تعیین شده، یک گراف نمونه برای کلمات منتخب ساخته می‌شود. در مرحله نهایی، میزان تلاقی خوشه‌های مختلف در گراف حاصل بررسی می‌شود.

در سال ۲۰۲۱، نویسندگان [۵۶] برای مقایسه روش‌های متفاوت آموزش بازنمایی توزیعی جملات، به مقایسه بازنمایی‌های حاصل می‌پردازند. ابتدا با استفاده از معیار تشابه فاصله اقلیدوسی[90]، هر بازنمایی به یک گراف نگاشت می‌شود. در این گراف، رئوس نماینده جملات و یال‌ها نشان دهنده کوچکتر بودن فاصله اقلیدوسی از یک آستانه از پیش تعیین شده است. پس از ساخت گراف جملات، کیفیت خوشه‌ها، اندازه قطر و توزیع درجات در گراف‌های حاصل مقایسه می‌شود.

در پژوهشی [۵۷] در سال ۲۰۲۱، نویسندگان تفاوت بازنمایی کلمات انتزاعی و کلمات عینی را با استفاده از شاخص‌های علم شبکه بررسی می‌کنند و مشاهده کردند که کلمات عینی جوامع متراکم بیشتری دارند.

[87] Biemann
[88] Motif Analysi
[89] Veremyev
[90] Euclidean Distance

۴.۲ تاریخچه ابزارها و روش‌های مقایسه شبکه‌های غیرهم‌اندازه

با توجه به رشد بسیار سریع استفاده از دادگان شبکه‌ای در حوزه‌های متفاوت، انتخاب روش‌ها و طراحی چهارچوب‌های مناسب به منظور مقایسه این شبکه‌ها اهمیت زیادی پیدا کرده است. بر این اساس، پیکاردی[۹۱] و همکاران [۵۸] در سال ۲۰۱۹ یک مطالعه مروری از انواع روش‌های موجود برای مقایسه شبکه‌ها انجام داده‌اند و راه‌کارهایی برای انتخاب روش مقایسه بسته به ماهیت شبکه‌های مورد مطالعه و هدف مقایسه آن‌ها ارائه می‌کنند. آن‌ها دو حالت متفاوت را برای مقایسه شبکه‌ها در نظر می‌گیرند؛ در حالت اول تناظر رئوس از قبل شناخته شده[۹۲] است اما در حالت دوم شناختی از تناظر میان رئوس نداریم[۹۳]. به طور دقیق‌تر، در حالت اول مجموعه رئوس در دو شبکه مورد مقایسه یکسان است و یا همپوشانی خوبی دارند و هم‌چنین تناظر یک به یک رئوس آن‌ها نیز از قبل شناخته شده است. در حالت دوم، هر دو شبکه فرضی فارغ از هم‌اندازه و یا متناظر بودن می‌توانند با یک‌دیگر مقایسه شوند. واضح است که که حالت اول در واقعیت کم‌تر رخ می‌دهد.

نویسندگان در مورد حالت دوم بیان می‌کنند که رویکردهای مرتبط با مقایسه گراف‌های غیرهم‌اندازه معمولا با اندازه‌گیری شاخص‌های سراسری در دو شبکه مورد مقایسه به تعریف یک معیار فاصله[۹۴] می‌رسند و با توجه به آن، میزان تشابه دو شبکه را بررسی می‌کنند. برای مثال، با اندازه‌گیری شاخص‌های ضریب خوشه‌بندی، قطر و یا متوسط فاصله، می‌توان دو شبکه غیرهم‌اندازه را مقایسه کرد. این در حالی است، که شبیه بودن مقادیر حاصل از این معیارها میان دو شبکه لزوما شباهت ساختاری آن‌ها را نشان نمی‌دهد [۵۸]. در این باره، کوستا[۹۵] و همکاران [۵۹] نشان می‌دهند که برخی شاخص‌های سراسری به اندازه شبکه حساس بوده و با زیاد شدن اندازه شبکه تغییر می‌کنند. آن‌ها هم‌چنین نشان می‌دهند که تغییرات مقادیر شاخص‌ها به ماهیت شبکه نیز بستگی دارد. برای مثال، آن‌ها سه شاخص متوسط فاصله، ضریب خوشه‌بندی سراسری و ضریب همسان‌گرایی درجه‌ای را برای مدل تصادفی اردوش-رنیی[۹۶] [۶۰]، مدل جهان‌کوچک واتز-اشتروگتز[۹۷] [۶۱] و مدل فارغ از مقیاس باراباشی-البرت[۹۸] [۶۲] اندازه گرفتند. آن‌ها نشان دادند که در دو مدل اردوش-رنیی و باراباشی-البرت، با افزایش اندازه شبکه، مقدار ضریب خوشه‌بندی زیاد می‌شود اما مقدار متوسط فاصله تغییر مشهودی نمی‌کند. این در حالی است که در

[91] Piccardi
[92] Known Node Correspondence
[93] Unknown Node Correspondence
[94] Distance
[95] Costa
[96] Erdos-Renyi Random Graphs
[97] Watts-Strogatz Small-world models
[98] Barabasi-Albert Scale-free Networks

مدل واتز–اشتروگتز، با افزایش اندازه شبکه، مقدار متوسط فاصله با شیب زیادی کاهش پیدا می‌کند اما ضریب خوشه‌بندی تغییر محسوسی نمی‌کند. به علاوه، با توجه به این‌که تنها مدلی که در آن اتصال ترجیحی وجود دارد مدل باراباشی–البرت است، افزایش اندازه شبکه تنها در این مدل موجب افزایش نسبی ضریب همسان‌گرایی درجه‌ای می‌شود.

شواهد این چنینی نشان می‌دهد برخی شاخص‌ها به اندازه شبکه حساس هستند و مقایسه مستقیم شبکه‌های غیرهم‌اندازه لزوما نتایج درستی درباره شباهت دو شبکه به دست نمی‌دهد. به علاوه، واضح است که مقایسه شبکه‌های غیرهم‌اندازه تنها با استفاده از شاخص‌های سراسری به ما اجازه بررسی شباهت‌های ساختاری در سطوح میانی و محلی را نمی‌دهد.

۵.۲ نتیجه‌گیری

در این بخش، به منظور شناسایی محدودیت‌های موجود در پیشینه تحقیق مطالعه ساختاری بازنمایی‌های معنایی، این موضوع از حوزه‌های مختلف بررسی شد. در ادامه جدولی حاوی مقالات مرتبط با موضوع تحقیق جاری آورده شده که زمینه‌های بهبود مطالعات پیشین را نشان می‌دهد. با توجه به جدول ارائه شده می‌بینیم که رویکردهای ارائه شده به منظور مطالعه بازنمایی‌های معنایی جامعیت و گوناگونی شبکه‌های معنایی را در نظر نگرفته‌اند. این در حالی است که کشف شباهت‌های موجود میان بازنمایی‌های معنایی رایج، می‌تواند درک بهتری از روش مدل‌سازی این بازنمایی‌ها به دست دهد. به علاوه، با مطالعه بازنمایی‌های معنایی می‌توان به فهم بهتری از چگونگی عملکرد هر مدل در امور پردازش معنا دست پیدا کرد.

از طرفی، با توجه به اینکه بازنمایی‌های معنایی مورد مطالعه در این پژوهش هم‌اندازه و یا متناظر نیستند، روش‌های موجود برای مقایسه شبکه‌های غیرهم‌اندازه، همان‌طور که بحث شد، لزوما نتایج قابل اعتمادی ارائه نمی‌کنند. با توجه به این موضوع، چهارچوب جدیدی مبتنی بر مقایسه بازنمایی‌های معنایی غیرهم‌اندازه معرفی خواهیم کرد که مقایسه میان شبکه‌های غیرهم‌اندازه را میسر کند. این چهارچوب می‌تواند در حوزه‌های دیگری

غیر از مطالعه معنا که از علم شبکه بهره می‌گیرند نیز به کار برده شود.

	مرجع	رویکرد	هدف	مزایا/نوآوری	معایب/محدودیت
مطالعه ساختاری بازنمایی معنایی و زبانی	۲۰۲۳ [۴۵]	استفاده از معیار ناهمگونی در فضای پیوسته	بررسی تاثیر معماری مدل بر بازنمایی نهایی اصوات کلمات زبان	گسترش کاربرد معیار ناهمگونی که پیش‌تر صرفا در فضای متنی استفاده شده بود.	
	۲۰۲۲ [۱۹]	استفاده از معیار ناهمگونی در فضای پیوسته	مقایسه بازنمایی‌های معنایی تک‌زبانه و چندزبانه	با استفاده از مطالعه بازنمایی معنایی چند زبانه، نشان می‌دهند که افزایش همگونی عملکرد آن‌ها را بهبود می‌بخشد.	تنها یک مدل ساخت بازنمایی مورد بررسی قرار گرفته است. مطالعه سایر بازنمایی‌های چند زبانه از جمله بازنمایی‌های انسانی، می‌تواند نتیجه‌گیری جامع‌تری به دست دهد.
	۲۰۲۲ [۴۴]	استفاده از معیار ناهم‌گونی در فضای پیوسته	مقایسه میزان حساسیت به بافت جملات در بازنمایی‌های مبتنی بر بافت جی‌پی‌تی، برت و المو و مقایسه میزان حساسیت به بافت در لایه‌های مختلف	پیشنهاد روشی به منظور بررسی کیفیت بازنمایی‌های پویا	به جهت جامعیت، مقایسه ناهمگونی می‌تواند میان بازنمایی‌های ایستا و پویا نیز انجام گیرد.

	مرجع	رویکرد	هدف	مزایا/نوآوری	معایب/محدودیت
	۲۰۲۲ [۱۸]	استفاده از روش‌های کاهش بعد در فضای پیوسته	شناسایی محورهای حساس به زبان و محورهای فارغ از زبان در بازنمایی چندزبانه	طرح یک چهارچوب مبتنی بر کاهش بعد که در شناسایی محورهای حساس به زبان در بازنمایی‌های چند زبانه موثر است.	تنها بازنمایی چندزبانه توزیعی مطالعه شده است. مطالعه بازنمایی چندزبانه انسانی می‌تواند مقایسه جامع‌تری فراهم کند.
	۲۰۲۱ [۵۷]	استفاده از شاخص‌های علم شبکه	مقایسه توزیع کلمات در بازنمایی کلمات انتزاعی و کلمات عینی	نشان داده‌اند کلمات عینی زبان ساختار متراکم‌تری از کلمات انتزاعی دارند.	مطالعه تنها در زبان فرانسه انجام شده و تعمیم‌پذیری آن مشخص نیست.
	۲۰۲۱ [۵۶]	استفاده از شاخص‌های علم شبکه	مطالعه و ارزیابی روش‌های مختلف تعبیه جملات	ساخت یک بازنمایی شبکه‌ای از فضای برداری مربوط به جملات و زیر جملات. مطالعات تطبیقی قبلی عمدتا فقط از کلمات، بازنمایی شبکه‌ای ساخته‌اند.	

	مرجع	رویکرد	هدف	مزایا/نوآوری	معایب/محدودیت
	۲۰۲۰ [۵۵]	استفاده از شاخص‌های علم شبکه	مطالعه چگونگی تشخیص کلمات چند معنایی در بازنمایی توزیعی برت	رویکرد نوین اتخاذ شده که با انتخاب یک نمونه از کلمات چندمعنایی و تک معنایی و با محاسبه شباهت کسینوسی میان آن‌ها گراف متناظر تولید شده و از هم‌پوشانی آن‌ها به منظور مطالعه چندمعنایی استفاده می‌شود.	این مطالعه می‌تواند درمورد دیگر مدل‌های توزیعی نیز انجام شود تا میزان تعمیم‌پذیری آن سنجیده شود.

	مرجع	رویکرد	هدف	مزایا/نوآوری	معایب/محدودیت
	۲۰۱۹ [۱۰]	استفاده از شاخص‌های علم شبکه	مقایسه بازنمایی‌های معنایی مبتنی بر دانش انسانی و بازنمایی‌های توزیعی معنا	برای اولین بار به یک مقایسه تطبیقی میان دو نوع بازنمایی مبتنی بر دانش انسانی و توزیعی پرداخته شده است.	بازنمایی‌هایی مورد مطالعه از جامعیت کافی برخوردار نیستند و تعمیم‌پذیری نتایج چندان قابل اعتماد نیست. از طرفی مقایسه مستقیم مقادیر حاصل از شاخص‌های سراسری مبنای مقایسه قرار گرفته که لزوما بازنمای شباهت میان گراف‌های معنایی غیرهم‌اندازه نیست.
	۲۰۱۸ [۱۵]	استفاده از معیار تمرکز و پراکندگی در فضای پیوسته	مطالعه تاثیر استفاده از روش‌های مختلف تعبیه گراف دانش بر بازنمایی پیوسته نهایی	معرفی دو معیار جدید برای مطالعه بازنمایی‌های معنایی	معیارهای معرفی شده جنبه‌های مختلف یک بازنمایی را مورد توجه قرار نمی‌دهند.

	مرجع	رویکرد	هدف	مزایا/نوآوری	معایب/محدودیت
	۲۰۱۷ [۱۷]	استفاده از معیار ناهمگونی در فضای پیوسته	به منظور بررسی خواص بازنمایی حاصل از مدل اسکیپ‌گرام معیار ناهمگونی را معرفی و استفاده می‌کند.	معرفی یک معیار جدید به منظور شناسایی الگوهایی از بازنمایی کلمات که پیش‌تر با استفاده از روش‌های کاهش بعد میسر نبود. نشان می‌دهد که اسکیپ‌گرام با وجود ایستا بودن ساختار بسیار ناهمگونی دارد.	
	۲۰۱۲ [۵۴]	استفاده از شاخص‌های علم شبکه	تشخیص جملات تولید شده توسط مدل زبانی n-گرام وجملات تولید شده توسط انسان‌ها	استفاده از شاخص موتیف برای تشخیص جملات تولیدی مدل زبانی آماری از جملات انسانی	شناسایی موتیف‌ها در بازنمایی‌های معنایی بزرگ هزینه‌بر است.

	مرجع	رویکرد	هدف	مزایا/نوآوری	معایب/محدودیت
مقایسه شبکه‌های غیرهم‌اندازه	۲۰۱۹ [۵۸]	مطالعه مروری بر روش‌های مقایسه ساختارهای شبکه‌ای	یک چهارچوب جامع از روش‌های موجود ارائه می‌شود. در این چهارچوب روش‌های مقایسه شبکه‌ها بسته به هم‌اندازه بودن یا نبودن آن‌ها به دو دسته کلی تقسیم می‌شود.	بازشناسی روش‌های متناسب برای مقایسه شبکه‌های غیرهم‌اندازه و محدودیت‌های آن‌ها. بیان می‌کند که شباهت نتایج حاصل از مقایسه شاخص‌های سراسری دو شبکه لزوما شباهت آن‌ها را تضمین نمیکند.	ضمن شناخت محدودیت واقع بر استفاده مستقیم از شاخص‌های سراسری برای مقایسه شبکه‌های غیرهم‌اندازه، درباره راهکار مرتبطی صحبت نمی‌شود.
	۲۰۰۷ [۵۹]	مطالعه مروری درباره انواع روش‌های مطالعه ساختارهای شبکه‌ای	ارائه چهارچوب مطالعه شبکه‌های پیچیده	نشان می‌دهند که برخی شاخص‌های سراسری، از جمله متوسط فاصله، ضریب خوشه بندی و گاهی ضریب همسانگرایی درجه‌ای مستقیما به تغییرات اندازه شبکه‌ها حساس هستند.	

فصل ۳

روش تحقیق

۱.۳ مقدمه

در سال‌های اخیر، معیارها و شاخص‌های علم شبکه به منظور مطالعه پدیده‌های طبیعی مختلفی کارآمد بوده‌اند. استفاده از این نوع بازنمایی از چند جهت حائز اهمیت است. نخست آنکه بازنمایی شبکه‌ای یک پدیده امکان مطالعه آن را از جنبه‌های مختلف و تازه‌ای را فراهم می‌کند. از جمله می‌توان به امکان مطالعه یک شبکه پیچیده در سطوح مختلف محلی، میانی و سراسری اشاره کرد که در هر سطح شاخص‌های مختلف هرکدام امکانات گوناگونی را ارائه میکنند. به علاوه، درصورت وجود بازنمایی‌های مختلف برای یک پدیده، به منظور مقایسه آنها، تبدیل همه بازنمایی‌ها به یک بازنمایی شبکه‌ای، مقایسه بازنمایی‌های اولیه را تسهیل می‌کند. مزیت دیگر این نوع بازنمایی، شهودی بودن نتایج حاصل از به‌کارگیری شاخص‌های علم شبکه نسبت به روش‌های دیگر مطالعه یک پدیده است. این موضوع، به خصوص در جهت مطالعه فضاهای برداری اهمیت بیشتری پیدا میکند. از آنجایی که بازنمایی معنایی حاصل از آموزش یک شبکه عصبی همواره به دلیل شفاف نبودن روند تصمیم‌گیری شان مورد انتقاد قرار گرفته‌اند، بازنمایی فضای توزیعی حاصل به صورت یک شبکه متصل از گره‌ها و یال‌ها، امکانات بیشتری به منظور مطالعه این فضاها به صورتی که برای انسان قابل درک باشد، فراهم می‌کند.

در این بخش به شرح چهارچوب پیشنهادی برای مطالعه و طبقه‌بندی بازنمایی‌های معنایی توزیعی و بازنمایی‌های معنایی متکی بر دانش انسانی میپردازیم. در اولین قدم، برای آشنایی خواننده با تعاریف پایه علم شبکه، این مفاهیم به همراه نمادهای مورد استفاده در بافت شبکه‌های معنایی معرفی می‌شوند. سپس، دادگان و مدل‌های معنایی مورد

استفاده در این پژوهش معرفی می‌شوند. در مرحله بعدی، مراحل نگاشت یک بازنمایی معنایی به گراف متناظر آن شرح داده می‌شود. با توجه به اینکه این نگاشت برای فضاهای معنایی متکی بر دانش انسانی و فضاهای توزیعی متفاوت است، مراحل هر یک به صورت جداگانه شرح داده خواهد شد.

پس از شرح چگونگی دست‌یابی به یک نوع بازنمایی مشترک شبکه‌ای برای تمامی بازنمایی‌های معنایی معرفی شده، به شرح چگونگی استفاده از شاخص‌ها و معیارهای مبتنی بر علم شبکه می‌پردازیم. شاخص‌های شبکه در دو سطح سراسری و میانی معرفی می‌شوند و چگونگی استفاده از آن‌ها در گراف‌های معنایی نیز توضیح داده می‌شود.

در مرحله بعدی، به طبقه‌بندی بازنمایی‌های معنایی با استفاده از خواص ساختاری آن‌ها می‌پردازیم. با توجه به غیرهم‌اندازه بودن گراف معنایی حاصل از هر یک از بازنمایی‌های معنایی، به منظور طبقه‌بندی آن‌ها یک رویکرد آماری جدید معرفی شده و مراحل آن شرح داده می‌شود.

در نهایت الگوریتم مورد استفاده جهت طبقه‌بندی بازنمایی‌ها نیز معرفی می‌شود.

۲.۳ مفاهیم پایه

تعریف ۱.۲.۳. یک گراف معنایی ساده و بدون جهت G که دارای V عدد گره و E عدد یال باشد، به صورت $\mathbf{G} = (V, E)$ تعریف می‌شود.

تعریف ۲.۲.۳. در یک گراف معنایی G، هر گره $v \in V$ می‌تواند نمایانگر یک تک کلمه (مثال: bad) و یا عبارت چندکلمه‌ای (مثال: too bad) باشد، به همین علت، از این پس برای داشتن یک تعریف ثابت برای گره‌ها اینطور میگوییم که هر گره درواقع نمایانگر یک مفهوم[1] در زبان طبیعی[2] است. مفاهیم مختلف در یک گراف معنایی به وسیله یک یال $e \in E$ به هم متصل می‌شوند که هر یال بیانگر وجود یک ارتباط معنایی[3] میان دو مفهوم است. معنا و تفسیر این ارتباط در فضاهای معنایی مختلف متفاوت است و هنگام معرفی هر فضای معنایی به تعریف دقیق‌تر یال در هر یک از این فضاها خواهیم پرداخت.

تعریف ۳.۲.۳. ماتریس مجاورت یک گراف با نماد A نمایش داده می‌شود و از آن به منظور نمایش گراف

[1] Concept
[2] Natural Language
[3] Semantic relation

استفاده می‌شود. در این ماتریس، اگر درایه سطر iام و ستون jام برابر یک باشد، این امر نشانگر وجود یال میان دو گره iام و jام و صفر بودن این درایه نشانگر عدم یال میان دو گره مذکور است.

۳.۳ نگاشت یک بازنمایی معنایی اولیه به بازنمایی شبکه‌ای متناظر آن

اولین قدم در جهت مقایسه بازنمایی‌های معنایی مختلف، دستیابی به یک نوع بازنمایی مشابه است. همانطور که در مقدمه فصل جاری گفته شد، روش نگاشت یک بازنمایی معنایی به همتای گرافی آن به نوع بازنمایی اولیه بستگی دارد. به‌علاوه، با توجه به توضیحات فصل دوم، یادآوری می‌شود که انواع بازنمایی‌های معنایی را می‌توان به دو دسته کلی انسانی (گسسته) و توزیعی (پیوسته) تفکیک کرد. در این بخش، هدف نگاشت هر دو دسته به یک نوع بازنمایی مشابه به منظور فراهم کردن امکان مقایسه آن‌ها می‌باشد.

در ادامه ابتدا به شرح روش نگاشت یک بازنمایی معنایی انسانی به گراف متناظر آن می‌پردازیم. سپس، مراحل نگاشت یک فضای معنایی توزیعی به گراف آن توضیح داده خواهد شد. هر نگاشت در اصل یک سیاست تخصیص یال است و تخصیص یال یا عدم آن بین دو مفهوم، بر اساس تعریف ارائه شده از رابطه معنایی در بازنمایی معنایی اولیه صورت می‌گیرد.

۱.۳.۳ نگاشت یک پایگاه داده واژگان به یک شبکه معنایی

در این پژوهش، سه پایگاه داده واژگان مختلف مورد استفاده قرار گرفته که در فصل دوم به طور مفصل به شرح ویژگی‌ها و ساختار هر یک از آنها پرداخته شده است. لذا در این بخش صرفا به توضیح کوتاهی درمورد تعریف رابطه معنایی در هر یک از آنها بسنده می‌کنیم و به شرح روش ساخت یک شبکه معنایی از هر یک از این پایگاه‌های داده خواهیم پرداخت.

۱.۱.۳.۳ نگاشت وردنت به یک بازنمایی شبکه‌ای

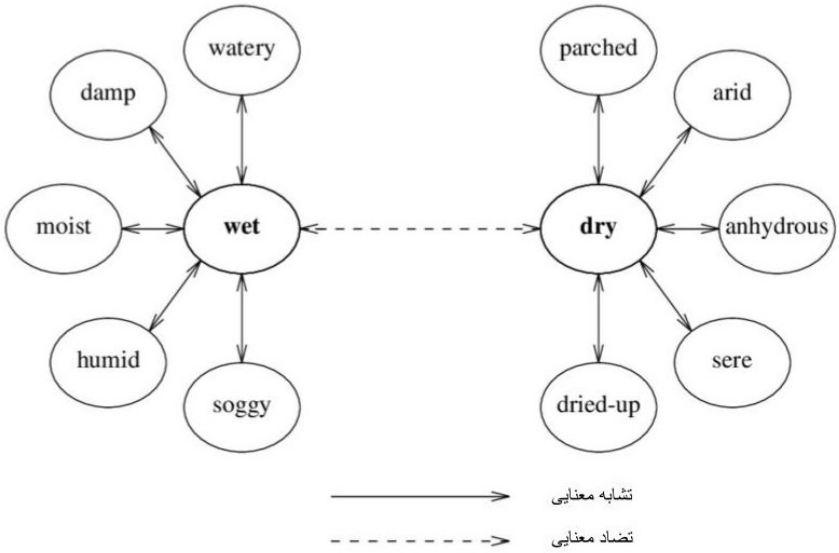
برای انجام این نگاشت ابتدا نیاز است که بدانیم روابط بین مفاهیم در این پایگاه داده به چه شکل تعریف شده است. در این پایگاه داده، واژگان به مجموعه‌هایی از مترادف‌های شناختی[۴] گروه‌بندی شده‌اند و هر دو واژه‌ای که در یک مجموعه ترادف حضور داشته باشند با یکدیگر رابطه معنایی دارند. از طرفی ممکن است هر یک از این مجموعه‌ها نیز با مجموعه‌های دیگر در رابطه معنایی قرار داشته باشند [۲۴].

در این پژوهش، مجموعه مورد نظر از کتابخانه NLTK[۵] در پایتون گردآوری شده است [۶۳]. این مجموعه داده در زمان انجام این پژوهش شامل حدودا ۱۱۷ هزار مجموعه ترادف بوده است که جمعا حدود ۱۴۸ هزار مفهوم را در برمی‌گرفته است.
با توجه به تعریف روابط در این پایگاه داده، سیاست تخصیص یال به این صورت خواهد بود که اگر دو کلمه حداقل یک بار با هم در یکی از مجموعه‌های هم‌معنایی مشاهده شوند، در گراف مقصد میان آن‌ها یک یال در نظر گرفته می‌شود. به بیان دیگر:

$$S = \{M_k | ۱ < k < n\}, \qquad A_{ij} = \begin{cases} ۱, & (\exists M_k | w_i \in M_k \wedge w_j \in M_k) \text{ اگر} \\ ۰, & \text{در غیر این صورت} \end{cases} \tag{۱.۳}$$

در رابطه بالا، S مجموعه گروه‌های ترادف ارائه شده در وردنت، n تعداد این گروه‌ها و M یک گروه ترادف فرضی در مجموعه S است. A نمایانگر ماتریس مجاورت گراف مقصد و w_i و w_j هر کدام یک کلمه در وردنت می‌باشند.

[4] Cognitive Synonyms
[5] Natural Language Processing Toolkit



شکل ۱.۳: بازنمایی معنایی وردنت [۱]

۲.۱.۳.۳ نگاشت فرهنگ‌نامه موبی به یک بازنمایی شبکه‌ای

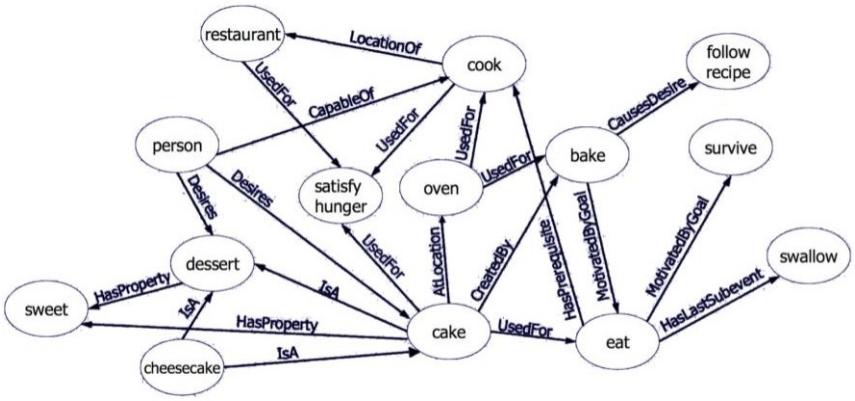
موبی، در واقع یک فرهنگ معنایی است که در آن به هر مدخل، تعدادی کلمه که از لحاظ معنایی به مدخل شباهت دارند نسبت داده شده است. به منظور نگاشت این فرهنگ معنایی به یک گراف متناظر با آن، سیاست تخصیص یال به این صورت است که اگر کلمه‌ای در مجموعه کلمات مرتبط با یک مدخل حضور داشته باشد، میان آن کلمه و مدخل مورد نظر یک یال برقرار می‌شود. در نسخه‌ای[۶] از موبی که در اینجا از آن استفاده کرده‌ایم، حدود سی و دو هزار مدخل و صدوهفت هزار کلمه یکتا یافت می‌شود.

۳.۱.۳.۳ نگاشت کانسپت‌نت به یک بازنمایی شبکه‌ای

ساختار اصلی کانسپت‌نت به گونه‌ای ارائه شده است که می‌توان آن را به خودی خود یک شبکه معنایی دانست. در این شبکه، گره‌ها نماینده مفاهیم و یال‌ها نمایانگر وجود و نوع رابطه میان دو مفهوم هستند. مجموعه داده‌ای که در این پژوهش مورد استفاده قرار گرفته[۷]، شامل تعداد زیادی یال از پیش تعریف شده است. به این صورت که هر رابطه میان دو مفهوم به صورت یک یال وزن‌دار ارائه شده است. علاوه بر وزن یک رابطه، نوع رابطه موجود

[6] http://onlinebooks.library.upenn.edu/webbin/gutbook/lookup?num=3202
[7] conceptnet.io

میان دو کلمه و همچنین زبان مرجع هر دو کلمه نیز از سری اطلاعات دیگری هستند که در این مجموعه داده ارائه شده‌اند. از آنجایی که این شبکه معنایی یک شبکه چندزبانه است و ممکن است روابط معنایی موجود تک زبانی و یا میان‌زبانی باشند، در این پژوهش از میان یال‌های ارائه شده تنها یال‌هایی استخراج شده‌اند که هر دو گره مرتبط با آن یال، مرجعشان زبان انگلیسی باشد. در بخش بحث و نتیجه‌گیری درباره روش‌های بهره‌گیری از فضاهای معنایی چندزبانه بیشتر صحبت خواهیم کرد اما در حال حاضر، تنها بخش انگلیسی زبان مورد تمرکز این پژوهش خواهد بود. برای نگاشت کانسپت‌نت به یک گراف، تنها لازم است در صورت تعریف یک یال در این مجموعه داده میان دو کلمه انگلیسی، یال متناظر با آن را در گراف مقصد لحاظ کنیم. گراف حاصل بدون جهت و بدون وزن خواهد بود. به علاوه، به منظور همگن‌سازی بازنمایی‌ها جهت مقایسه، نوع رابطه معنایی در یال‌های موجود در این پژوهش لحاظ نشده‌اند.



شکل ۲.۳: بازنمایی معنایی کانسپت‌نت [۲]

۲.۳.۳ نگاشت یک فضای معنایی ازپیش‌آموزش‌دیده به یک شبکه معنایی

در حال حاضر، یادگیری یک بازنمایی معنایی با استفاده از روش‌های یادگیری ماشین یک امر بسیار رایج است. به این منظور، برای دست‌یابی به یک فضای معنایی توزیعی از روش‌های بسیار گسترده‌ای می‌توان بهره گرفت که در فصل دوم به طور مفصل در این باره صحبت شد. به طور کلی، می‌توان گفت که در این روش یک مدل زبانی با استفاده از یک مجموعه متنی نسبتا بزرگ آموزش داده شده و با توجه به ساختار مدل مذکور، نتیجه این آموزش دستیابی به یک توزیع برای کلمات زبان است. یکی از نقاط قوت این روش، امکان اضافه کردن اطلاعات دیگری علاوه بر داده متنی در زمان آموزش مدل است. از جمله می‌توان به اطلاعات موجود در گراف‌های دانش،

اطلاعات تصویری و صوتی و یا حتی توزیع کلمات در زبان‌های دیگر اشاره کرد. به همین علت در این پژوهش سعی شده است از گروه‌های مختلف فضاهای معنایی برداری یک نماینده از هر گروه به منظور مطالعه و مقایسه در نظر گرفته شود. در ادامه ابتدا به توضیح کوتاهی جهت معرفی هر یک از فضاهای مورد مطالعه می‌پردازیم و سپس به توضیح مراحل نگاشت آن‌ها به بازنمایی شبکه‌ای متناظرشان خواهیم پرداخت. (مجموعه فضاهای معنایی مورد مطالعه در این تحقیق، ساختار مدل آن‌ها و همچنین داده مورد استفاده برای آموزش هرکدام از آن‌ها جهت دستیابی به یک بازنمایی معنایی برداری در فصل دوم به طور کامل معرفی شدند، لذا در صورت نیاز به توضیح بیشتر به فصل دوم مراجعه شود.)

۱.۲.۳.۳ فضاهای معنایی ازپیش‌آموزش‌دیده

با توجه به دسته‌بندی این گروه از بازنمایی‌های معنایی، فضاهای معنایی ازپیش‌آموزش‌دیده مورد مطالعه در این پژوهش حاصل آموزش مدل‌های زیر هستند:

مدل Word2Vec فضای معنایی چندبعدی حاصل از آموزش مدل کیسه لغات پیوسته[8] با استفاده از پیکره خبری گوگل[9]. فضای حاصل شامل مجموعه واژگانی با اندازه سه میلیون کلمه است و درواقع نگاشتی است یک به یک، از کلمات به بردار متناظرشان.

مدل BERT2Static فضای معنایی در این مورد، حاصل آموزش یک مدل الهام‌گرفته از معماری کیسه لغات پیوسته به نام Sent2Vec [۶۴]، بر روی پیکره ویکیپدیای انگلیسی[10][11] می‌باشد. در حین آموزش این مدل در واحد رمزگذار بافتار[12]، به هدف بهره‌گیری از تعبیه واژگانی پویا[13] در مدل زبانی برت، بردار کلمات بافت[14] واژه مخفی[15] نه از مدل Word2Vec، بلکه از لایه آخر مدل برت گرفته شده‌اند. در نتیجه فضای معنایی حاصل از هر دو نوع

[8] Continuous Bag of Words
[9] Google News
[10] English Wikipedia Dump
[11] https://huggingface.co/datasets/wikipedia
[12] Context Encoder Unit
[13] Dynamic Word Embedding
[14] Context Words
[15] Masked Word

تعبیه کلمات ایستا[۱۶] و پویا بهره می‌برد.

مدل VisualWord2Vec این مدل، در واقع روشی برای لحاظ کردن اطلاعات دیداری در یک فضای برداری آموزش‌دیده از داده متنی است. در این روش، یک مدل شبکه عصبی با آموزش بر روی مجموعه‌ای از جفت نمونه‌های تصویر-متن، روابط معنایی که در داده تک‌وجهی متنی لزوما قابل دسترسی نیستند را فرامی‌گیرد. در حین آموزش این مدل، مقداردهی اولیه وزن‌های لایه اول نه به صورت تصادفی بلکه با وزن‌های به دست آمده از آموزش مدل کیسه لغات پیوسته بر روی مجموعه داده متنی ویکیپدیای انگلیسی انجام می‌شود. بردارهای حاصل را می‌توان بهبود یافته بردارهای مدل Word2Vec با استفاده از روابط معنایی مبتنی بر اطلاعات تصویری تلقی کرد.

مدل Conceptnet Numberbatch آخرین فضای معنایی که در مطالعه ما مورد استفاده قرار می‌گیرد، فضای برداری Conceptnet Numberbatch می‌باشد که در تحقیق جاری صرفا از بخش انگلیسی آن استفاده خواهیم کرد. این بازنمایی برداری درواقع یک فضای معنای چندزبانه است و از ترکیب بردارهای Word2Vec، بردارهای GloVe، و دو داده واژگانی ساختارمند کانسپت‌نت و پیکره دگربیان[۱۷] [۶۵]به دست آمده است. برای ساخت فضای نهایی، از مدل ازپیش‌آموزش‌دیده Word2Vec بر روی مجموعه داده خبری گوگل استفاده شده است. آموزش مدل GloVe نیز از قبل روی داده کامن کرال[۱۸] [۱۹] انجام شده است.

شایان ذکر است که در این پژوهش، برای هر یک از فضاهای در نظر گرفته شده، هیچگونه آموزشی انجام نمی‌شود و تنها بردارهای ازپیش‌آموزش‌دیده هر یک از آنها بارگیری شده و به بازنمایی شبکه‌ای متناظرشان نگاشت می‌شوند. منابع تهیه هر یک از مدل‌ها نیز ارائه شده است.

[16] Static Word Embedding
[17] Paraphrase Database (PPDB)
[18] Common Crawl
[19] https://commoncrawl.org/

۲.۲.۳.۳ نگاشت فضای معنایی ازپیش‌آموزش‌دیده به گراف متناظر آن

همانطور که گفته شد، به دلیل برداری بودن این فضاهای معنایی، مراحل نگاشت آنها به یک شبکه معنایی نیز متفاوت از مراحل ذکر شده برای پایگاه داده‌های واژگان که در بخش پیش ارائه شد خواهد بود. اما این فرآیند برای تمامی فضاهای برداری یکسان خواهد بود و توضیح پیش رو درباره نگاشت هر یک از فضاهای برداری به بازنمایی شبکه‌ای متناظر آن صادق و یکسان است.

برای نگاشت یک فضای معنایی برداری به یک بازنمایی شبکه‌ای نیازمند یک سیاست تخصیص یال هستیم که بتواند مفاهیم مرتبط را در گراف مقصد به یکدیگر متصل کند و همچنین یالی میان دو مفهوم معنایی غیر مرتبط اختصاص ندهد. به علت پیوسته بودن نمایش مفاهیم در این دسته از فضاهای معنایی، تعیین سیاست تخصیص یال به‌سادگی آنچه که درباره‌ی نگاشت پایگاه‌های داده واژگانی مشاهده کردیم نخواهد بود.

همانطور که در فصل دوم به توضیح فضاهای برداری معنایی پرداخته شد، می‌دانیم که در چنین بازنمایی، معمولا یک کلمه به یک بردار مشخص در یک فضای K بعدی نگاشت می‌شود و برداری با اندازه K نمایانگر هر کلمه خواهد بود. در این فضای معنایی، هر چه بردار اختصاص‌یافته به دو کلمه فرضی به یکدیگر نزدیک‌تر باشند، احتمالا آن دو کلمه از لحاظ معنایی به یکدیگر شبیه‌تر هستند. با این توضیح، [۱۰] برای محاسبه میزان ارتباط معنایی دو کلمه در یک فضای چند بعدی، از مفهوم شباهت معنایی[20] استفاده میکنند. در اینجا این نکته قابل ذکر است که دو مفهوم شباهت معنایی و ارتباط معنایی[21] لزوما با هم یکسان نیستند [۶۶] اما به نظر می‌رسد که [۱۰] به منظور نگاشت فضای برداری به یک بازنمایی شبکه‌ای این تفاوت را در نظر نگرفته‌اند. به همین ترتیب ما نیز از همین فرض استفاده خواهیم کرد.

راه حل پیشنهادی پژوهش نام‌برده برای محاسبه شباهت معنایی میان دو کلمه، اندازه‌گیری میزان شباهت کسینوسی میان دو بردار نمایانگر دو کلمه مورد نظر در فضای معنایی است. به این صورت که اگر بردار دو کلمه w_i و w_j را به ترتیب v^i و v^j در نظر بگیریم، شباهت این دو کلمه با استفاده از رابطه زیر محاسبه خواهد شد:

$$sim(w_i, w_j) = \frac{\sum_{k=1}^{K} v_k^i v_k^j}{\sqrt{\sum_{k=1}^{K} (v_k^i)^2} \sqrt{\sum_{k=1}^{K} (v_k^j)^2}} \qquad (۲.۳)$$

[20] Semantic Similarity
[21] Semantic Relatedness

در مرحله بعدی، با در نظر گرفتن یک آستانه تشابه[۲۲] δ به صورت قراردادی، سیاست تخصیص یال به این صورت خواهد بود که اگر تشابه کسینوسی دو کلمه از آستانه مورد نظر بیشتر باشد، در گراف مقصد میان این دو مفهوم یک یال در نظر گرفته می‌شود:

$$A_{ij} = \begin{cases} ۱, & (sim(w_i, w_j) \geq \delta) \text{ اگر} \\ ۰, & \text{در غیر این صورت} \end{cases} \quad (۳.۳)$$

در قسمت بعدی به شرح چهارچوب پیشنهادی مقایسه بازنمایی‌های به دست آمده خواهیم پرداخت.

۴.۳ چهارچوب پیشنهادی جهت مطالعه و طبقه‌بندی بازنمایی‌های معنایی انسانی و توزیعی

پس از نگاشت همه بازنمایی‌های معنایی مورد مطالعه مطابق با توضیحات بخش پیشین به شبکه متناظر آن‌ها، حال می‌توانیم به مطالعه و طبقه‌بندی گراف‌های حاصل بپردازیم. در این قسمت، قبل از هر چیز ذکر این نکته ضروری است که انتخاب معیارهای مناسب برای مطالعه یک ساختار شبکه‌ای می‌تواند به چند علت امری چالش‌برانگیز باشد. نخستین مسئله، فراوانی و گوناگونی انواع شاخص‌ها و ابزارهای علم شبکه است چرا که ممکن است این فراوانی باعث سردرگمی پژوهشگرانی شود که به دنبال معیار مناسب برای مسئله خود می‌گردند. از طرفی دیگر، با وجود فراوانی این معیارها، همپوشانی میان آنها بعید نیست و لزوما نتایج کاملا مجزایی ارائه نمی‌کنند، با این وجود، از آنجایی که هر یک از معیارها تنها جنبه‌ای از خصوصیات یک شبکه را مورد توجه قرار می‌دهند، به کارگیری تعداد محدودی معیار علم شبکه به منظور مطالعه یک پدیده شبکه‌ای لزوما نتایج قابل اعتمادی به دست نمی‌دهد.

یکی دیگر از مشکلاتی که به طور مشخص در مورد پژوهش جاری صدق می‌کند، مسئله مقایسه ساختارهای شبکه‌ای است که از تعداد گره و یال یکسان برخوردار نیستند. این مسئله علی‌الخصوص زمان استفاده از معیارهایی که وابسته به اندازه شبکه هستند نمود بیشتری پیدا می‌کند.

[22] Similarity Threshold

در آخر ذکر این نکته ضروری است که با توجه به ماهیت زبان طبیعی و وجود تعداد زیاد مفاهیم و روابط میان آنها، تقریبا تمامی بازنمایی‌های معنایی مورد استفاده را می‌توان یک شبکه بزرگ [۲۳] [۶۷] تلقی کرد. مطالعه این نوع ساختارها ممکن است محدودیت‌هایی از جهت هزینه‌بر بودن محاسبات ایجاد کند.

با توجه به تمامی مسائل ذکر شده، معرفی یک چهارچوب مقایسه شبکه‌های معنایی که تا حد امکان مسائل ذکر شده را لحاظ کرده باشد، اهمیت پیدا می‌کند.

چهارچوب پیشنهادی این پژوهش را می‌توان به سه دسته کلی تقسیم کرد. دسته اول معیارهای مورد استفاده، عمدتا ویژگی‌های سراسری شبکه‌ها را مورد توجه قرار می‌دهند. در دسته دوم معیارها، به بررسی ویژگی‌های گراف‌ها در سطح میانی، از جمله مطالعه جوامع می‌پردازیم. دسته سوم، مبتنی بر مقایسه آماری میان گراف‌های معنایی و مدل‌های تصادفی متناظرشان است که امکان مقایسه گراف‌های غیرهم‌اندازه معنایی را به دست می‌دهد و از این رو امکان طبقه‌بندی گراف حاصل از بازنمایی‌های معنایی را نیز فراهم می‌کند.

۱.۴.۳ شاخص‌های سراسری

همان‌گونه که از نام این دسته از شاخص‌ها برمی‌آید، شاخص‌های سراسری مطالعه ساختارهای شبکه‌ای، عموما به بررسی جنبه‌های مختلف در کل ساختار می‌پردازند. به این معنا که در این مقیاس از مطالعه، از بررسی گره‌ها[۲۴] و جوامع[۲۵] فراتر رفته و به مطالعه ویژگی‌های برآمده از تمامی ساختار یک شبکه می‌پردازیم. ویژگی‌های ساختاری یک شبکه می‌تواند در درک ویژگی‌های رفتاری آن مورد استفاده قرار گیرد [۵۹]. در ادامه، معیارهای سراسری مورد استفاده در این پژوهش به منظور مطالعه ساختاری[۲۶] بازنمایی‌های معنایی، معرفی و بررسی خواهند شد.

۱.۱.۴.۳ شاخص‌های سراسری ساختاری

تراکم یال شاخص تراکم یال امکان اندازه‌گیری تراکم یال‌های موجود میان گره‌های یک شبکه را به دست میدهد. به بیانی دیگر، این شاخص امکان مقایسه میزان ارتباطات موجود در یک شبکه را به میزان ارتباطاتی که آن شبکه

[23] Big Network
[24] Nodes
[25] Communities
[26] Topological Characterization

قادر است در خود جا بدهد را به دست می‌دهد. نحوه محاسبه شاخص تراکم یال (ξ) در زیر آمده است. در این رابطه N و M به ترتیب تعداد گره‌های گراف و تعداد یال‌ها را نشان می‌دهد.

$$\xi = \frac{۲M}{N(N-۱)} \qquad (۴.۳)$$

میزان تراکم یال در یک بازنمایی معنایی، نشان می‌دهد که روابط معنایی بالقوه در آن بازنمایی لحاظ شده‌اند. بازنمایی چگال‌تر میل بیشتری به دربرگیری تعداد بیشتری رابطه معنایی از خود نشان می‌دهد و یک بازنمایی تنک[۲۷] ممکن است روابط معنایی مختلفی را در نظر نگرفته باشد.

شایان ذکر است که چگالی و تنک بودن یک شبکه از مفهوم بردارهای چگال و تنک معنایی متفاوت است. در مورد اول، چگالی بیشتر به معنای در نظر گرفتن روابط بیشتر میان مفاهیم در یک بازنمایی است، اما در مورد دوم، چگال بودن یک بردار تنها به روش بازنمایی کلمات ارتباط دارد و لزوما ارتباطی به روابط میان کلمات در آن بازنمایی ندارد.

متوسط درجه این معیار نمایانگر متوسط تعداد ارتباطات هر گره با گره‌های دیگر در یک شبکه است. در شبکه‌های معنایی، این شاخص نشان می‌دهد که هر مفهوم به طور متوسط چند ارتباط معنایی با مفاهیم دیگر دارد.

قطر در یک شبکه، قطر به طولانی‌ترین مسیری که ممکن است میان دو گره از آن شبکه موجود باشد اطلاق می‌شود.

متوسط کوتاه‌ترین مسیر پس از محاسبه کوتاه‌ترین مسیرها میان هر دو گره موجود در یک شبکه، میانگین طول همه این مسیرها، متوسط تعداد گره‌ای که باید از آن بگذریم تا از یک گره به گره‌ای دیگر در یک شبکه برسیم به دست می‌آید.

[27] Sparse

این شاخص در کنار شاخص قطر جهت مقایسه زمانی که برای رسیدن اطلاعات از یک گره شبکه معنایی تا گره‌های دیگر آن لازم است مورد استفاده قرار می‌گیرد.

شاخص ضریب خوشه‌بندی محلی این شاخص نشان میدهد که همسایگان یک گره مشخص در شبکه تا چه حد به هم متصل هستند. اگر k_i و L_i به ترتیب نمایانگر درجه گره i و تعداد اتصالات میان همسایگان این گره باشند، آنگاه ضریب خوشه‌بندی گره iام ($\mathbf{C_i}$) به صورت زیر تعریف می‌شود:

$$\mathbf{C_i} = \frac{۲L_i}{k_i(k_i - ۱)} \tag{۵.۳}$$

میزان خوشه‌بندی در سطح کل شبکه به واسطه میانگین ضریب خوشه‌بندی به ازای تمام گره‌ها محاسبه میشود. این معیار درواقع نشان می‌دهد چقدر احتمال دارد در صورت انتخاب تصادفی یک گره از شبکه، همسایگان آن به یک‌دیگر متصل باشند. متوسط ضریب خوشه‌بندی محلی ($\mathbf{C}$) به صورت زیر محاسبه می‌شود:

$$\mathbf{C} = \frac{۱}{N}\sum_{i=۱}^{N} C_i \tag{۶.۳}$$

شاخص ضریب خوشه‌بندی سراسری علاوه بر محاسبه میانگین ضریب خوشه‌بندی محلی برای تمام گره‌ها، راه دیگر محاسبه میزان خوشه‌بندی در سراسر یک شبکه، استفاده از شاخص ضریب خوشه‌بندی سراسری ($\tilde{\mathbf{C}}$) است. این شاخص در واقع نمایانگر نسبت سه‌تایی‌های بسته[۲۸] به تمام سه‌تایی‌های موجود در شبکه (باز یا بسته) می‌باشد و برای محاسبه آن داریم:

$$\tilde{\mathbf{C}} = \frac{\sum_{i,j,k} A_{ij}A_{jk}A_{ki}}{\sum_i k_i(k_i - ۱)} \tag{۷.۳}$$

نکته قابل توجه این است که شاخص خوشه‌بندی سراسری و متوسط شاخص خوشه‌بندی محلی لزوما با یکدیگر برابر نمی‌باشند. به علاوه، در تعریف ضریب خوشه‌بندی سراسری وزن بیشتری به گره‌های با درجات بالاتر داده می‌شود اما در شاخص دیگر، گره‌های با درجه پایین بیشتر مورد توجه قرار می‌گیرند [۶۸].

[28] Closed Triple (Triangle)

شاخص ضریب همسان‌گرایی درجه‌ای برای محاسبه میزان تمایل گره‌های هم‌درجه به اتصال با یکدیگر در یک شبکه، از ضریب همسانگرایی درجه‌ای استفاده می‌شود. موجودیت‌های یک شبکه ممکن است همسانگرا[۲۹]، خنثی[۳۰] و یا مخالف‌گرا[۳۱] باشد [۶۹]. این شاخص درواقع ضریب همبستگی پیرسون[۳۲] میان درجات گره‌های دو سر یک یال می‌باشد و به ازای تمامی یال‌ها در شبکه محاسبه می‌شود.

توزیع درجات گراف‌های معنایی به منظور مقایسه نحوه توزیع درجات در گراف‌های معنایی مورد مطالعه، توزیع آن‌ها رسم می‌شود. سپس، احتمال متناسب بودن توزیع‌های رایج درجات شبکه‌های پیچیده برای گراف‌ها بررسی می‌شود. از جمله این توزیع‌ها می‌توان به توزیع توانی[۳۳]، نمایی[۳۴] و نمایی کشیده[۳۵] اشاره کرد.

۲.۱.۴.۳ شاخص‌های مرکزیت

فراوانی استفاده از کلمات متفاوت در یک زبان می‌تواند بسیار مختلف باشد، برخی کلمات بسیار رایج هستند و برخی به ندرت مورد استفاده قرار می‌گیرند. [۷۰] ادعا می‌کنند که در یک زبان، کلمات هم‌معنای بیشتری برای مفاهیمی که با فراوانی بیشتری مورد استفاده قرار می‌گیرند یافت می‌شود. به بیان دیگر، اگر کلمه‌ای زیاد مورد استفاده قرار بگیرد، احتمال وجود کلمات هم‌معنا برای آن کلمه بالاتر است. اگر به نحوه ساخت بازنمایی‌های معنایی برداری نیز توجه کنیم، می‌بینیم که رخداد بیشتر یک کلمه، منجر به هم‌آیی بیشتر آن با سایر کلمات شده و در نتیجه انتظار می‌رود در بازنمایی نهایی برداری، بردار آن به بردار کلمات متعددی نزدیک باشد.

در این پژوهش، به منظور سنجش این پدیده از شاخص‌های مرکزیت استفاده می‌کنیم. به این منظور، مجموعه کلمات انگلیسی مرتب‌شده بر حسب فراوانی[۳۶] که از پیکره کلمات میلیاردی گوگل[۳۷][۳۸] به دست آمده، مورد استفاده قرار گرفته است.

[29] Assortative
[30] Neutral
[31] Disassortative
[32] Pearson Correlation Coefficient
[33] Power Law
[34] Exponential
[35] Stretched Exponential
[36] https://www.kaggle.com/datasets/rtatman/english-word-frequency?
[37] Google Web Trillion Word Corpus
[38] https://norvig.com/ngrams/

با توجه به توضیح داده شده، از آنجایی که انتظار می‌رود کلمات پرتکرار ارتباطات معنایی بیشتری داشته باشند، در هر یک از شبکه‌های معنایی ساخته شده، با بررسی تاثیرگذارترین رئوس، همپوشانی آن‌ها با کلمات پرتکرار زبان انگلیسی بررسی می‌شود.

با توجه به اینکه شاخص‌های مرکزیت متفاوت لزوما نتایج یکسانی به دست نمی‌دهند، به منظور اطمینان از اعتبار نتایج این قسمت از دو شاخص مرکزیت استفاده شده است.

شاخص مرکزیت بر حسب درجه اولین شاخص مرکزیت مورد استفاده، بر حسب درجه رئوس است. در این شاخص، هر چه یک راس درجه بالاتری داشته باشد، امتیاز مرکزیت بالاتری دریافت می‌کند.

شاخص مرکزیت برحسب الگوریتم پیج‌رنک شاخص دومی که مورد استفاده قرار گرفته است، شاخص پیج‌رنک[۳۹] می‌باشد. این معیار سراسری، به منظور محاسبه میزان تاثیرگذاری یک راس، علاوه بر شمارش تعداد اتصالات آن، میزان تاثیرگذاری رئوس همسایه آن راس را نیز مورد توجه قرار می‌دهد. در واقع در اینجا، میزان تاثیرگذاری هر راس، تابعی از تعداد همسایه‌های آن و میزان تاثیرگذار بودن هر یک از همسایه‌های آن است.

۲.۴.۳ طبقه‌بندی بازنمایی‌های معنایی بر اساس ویژگی‌های ساختاری سراسری

برای مقایسه ساختارهای شبکه‌ای با یک‌دیگر، همان‌گونه که گفته شد، ابزارها و روش‌های متعددی وجود دارد. با این وجود، اگر گراف‌هایی که قصد مقایسه آن‌ها را داریم در تعداد رئوس و یال یکسان و یا مشابه نباشند، آنگاه تحلیل نتایج حاصل از شاخص‌ها دشوار می‌شود. از آنجایی که شبکه‌های معنایی مورد مطالعه در این پژوهش اندازه‌های کاملا متفاوتی دارند، نیاز به یک چهارچوب مشخص که امکان مقایسه آن‌ها را به دست دهد ایجاد می‌شود. در این بخش به ارائه چهارچوبی جدید برای مقایسه شبکه‌های معنایی با اندازه‌های متفاوت می‌پردازیم. این چهارچوب می‌تواند در هر مسئله دیگری که مربوط به مقایسه شبکه‌های غیرهم‌اندازه باشد نیز، به کار گرفته شود.

در این چهارچوب ابتدا برای هر گراف معنایی مورد مطالعه نمونه مناسبی از گراف‌های تصادفی مشابه آن تولید می‌شود. سپس، شاخص‌های سراسری شبکه برای گراف معنایی و نمونه‌های تصادفی آن محاسبه می‌شوند. در این

[39] Pagerank Centrality

مرحله، با استفاده از رویکرد آماری که در ادامه معرفی می‌شود، به طبقه‌بندی بازنمایی‌های معنایی می‌پردازیم.

پیاده‌سازی این چهارچوب شامل چند مرحله است که در ادامه به تفصیل شرح داده می‌شود.

۱.۲.۴.۳ ساخت مدل پیکربندی تصادفی از یک گراف معنایی

در اولین مرحله از این چهارچوب، برای هر تعداد گرافی که به منظور مقایسه در اختیار داریم، تعداد مشخصی مدل پیکربندی تصادفی تولید می‌شود.

مدل پیکربندی تصادفی درواقع شبه‌گرافی[۴۰] است که توزیع درجات یکسانی با گراف اصلی دارد اما اتصالات آن به صورت تصادفی برقرار شده‌اند. تعداد مدل‌های تصادفی مورد نیاز برای مقایسه به کاربرد، اهداف و منابع مطالعه بستگی دارد. در این پژوهش، به منظور استفاده بهینه از منابع محاسباتی موجود، برای هر گراف معنایی، ده مدل پیکربندی تولید شده است. به بیان دیگر، جمعا دویست و سی مدل تصادفی تولید و استفاده شد. روش‌های متفاوتی برای ساخت یک مدل پیکربندی وجود دارد که پژوهشگر می‌تواند بسته به مقتضیات تحقیق خود یکی را برگزیند. در این پژوهش، با توجه به بی‌جهت بودن گراف‌ها و نبود ترجیح خاص در نتایج تولید مدل‌های تصادفی، برای تولید این مدل‌ها از پیاده‌سازی موجود در کتابخانه Networkx[۴۱] استفاده شده است. در این پیاده‌سازی، برای ساخت یک مدل پیکربندی تصادفی از یک گراف بدون جهت، ابتدا مجموعه درجات گراف مورد نظر به مجموعه ته‌یال‌های آن (نیمه‌یال)[۴۲] نگاشت می‌شود. در بستر مدل‌های پیکربندی، ته‌یال به انتهای تکرارشونده یال‌ها گفته می‌شود که برای تولید یک مدل تصادفی به کار می‌روند. ته‌یال‌ها با تکرار هر راس به تعداد درجه‌شان در گراف اصلی تولید می‌شوند. پس از تولید مجموعه ته‌یال‌ها، با تغییر تصادفی ترتیب آن‌ها و ترکیب جفت ته‌یال‌ها با هم، مدل پیکربندی تصادفی متناظر با گراف اصلی تولید می‌شود. واضح است که تعداد رئوس، یال‌ها و توزیع درجات مدل تصادفی مشابه گراف اصلی است.

الگوریتم مربوط به این پیاده‌سازی از کتابخانه NetworkX در ۱.۳ آورده شده است.

[40] Pseudograph
[41] `https://networkx.org/documentation/stable/reference/generators.html`
[42] Stub

الگوریتم ۱.۳ الگوریتم ساخت مدل پیکربندی تصادفی برای یک گراف

ورودی: توزیع درجات گراف G به عنوان مبنای تولید مدل پیکربندی تصادفی
خروجی: مدل پیکربندی تصادفی گراف

۱: تشکیل یک گراف تهی به عنوان مدل پیکربندی تصادفی
۲: تبدیل توزیع درجات ورودی به نیمه-یال‌ها
۳: تغییر ترتیب نیمه-یال‌ها به صورت تصادفی
۴: تشکیل مجموعه یال‌های جدید با ترکیب نیمه-یال‌ها
۵: ساخت مدل پیکربندی تصادفی با اضافه کردن یال‌های حاصل به گراف تهی ایجاد شده
۶: **بازگردان** مدل پیکربندی تصادفی

۲.۲.۴.۳ مقایسه گراف معنایی با مدل‌های پیکربندی تصادفی متناظر آن

پس از تولید مدل‌های پیکربندی تصادفی برای هر یک از گراف‌های معنایی مورد مطالعه، شاخص‌های سراسری که قصد مقایسه نتایج آن‌ها را داریم روی هر گراف معنایی و مدل‌های تصادفی آن اعمال می‌شوند. در این قسمت، شاخص‌های متوسط فاصله، ضریب خوشگی سراسری و محلی و همسان‌گرایی درجه‌ای استفاده قرار می‌گیرند. با داشتن نتایج چهار شاخص نام‌برده شده ابتدا به مقایسه هر گراف معنایی با مدل‌های تصادفی خودش می‌پردازیم. به این منظور، مقدار حاصل برای یک شاخص میان گراف معنایی و ده مدل تصادفی همتای آن با استفاده از آزمون آماری t تک‌نمونه‌ای[۴۳] مقایسه می‌شوند. به طور کلی این آزمون برای بررسی وجود و یا عدم تفاوت میان میانگین دو گروه نمونه به کار می‌رود. در این پژوهش، از آنجایی که یک گروه نمونه از مدل‌های تصادفی با توزیع نرمال برای تمامی شاخص‌ها و تنها یک مقدار برای گراف معنایی اصلی داریم، از آزمون آماری t تک‌نمونه‌ای استفاده شده است. این تعریف از این آزمون آماری امکان مقایسه یک مقدار مشخص با میانگین یک مجموعه نمونه به دست می‌دهد و به صورت زیر تعریف می‌شود:

$$\mathbf{t} = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \qquad (۸.۳)$$

[43] One-sample T-test

در این رابطه، مقدار t از تقسیم تفاوت میانگین مقادیر یک شاخص برای مدل‌های تصادفی و مقدار به دست آمده برای گراف اصلی به مقدار خطای استاندارد[۴۴] مقادیر آن شاخص برای مدل‌های پیکربندی تصادفی به دست می‌آید[۴۵].

پس از انجام مراحل گفته شده، یک ماتریس حاوی مقادیر حاصل از آزمون t به دست می‌آید. هر سطر این ماتریس نشانگر یکی از بیست و سه گراف معنایی مورد مطالعه است و هر ستون آن یکی از پنج شاخص شبکه را نشان می‌دهد. تمامی مراحل گفته شده در الگوریتم ۲.۳ به صورت خلاصه آمده است.

الگوریتم ۲.۳ الگوریتم مقایسه گراف معنایی با مدل‌های پیکربندی تصادفی آن

ورودی: A به عنوان مجموعه شامل n عدد گراف معنایی
ورودی: B به عنوان مجموعه شامل m شاخص شبکه
ورودی: c به عنوان تعداد مدل‌های پیکربندی تصادفی مطلوب برای هر گراف معنایی
خروجی: ماتریس با n سطر و m ستون حاوی مقادیر آزمون آماری t

۱: **برای** تمام گراف‌ها در مجموعه A در بازه ۱ تا n **انجام بده**
۲: **برای** مقادیر در بازه ۱ تا c **انجام بده**
۳: ساخت مدل پیکربندی تصادفی گراف با استفاده از توزیع درجات آن
۴: ذخیره مدل پیکربندی تصادفی
۵: **پایان حلقهٔ برای**
۶: **برای** تمام شاخص‌ها در مجموعه B در بازه ۱ تا m **انجام بده**
۷: محاسبه مقدار شاخص برای گراف اصلی
۸: محاسبه میانگین مقدار شاخص به ازای تمام مدل‌های پیکربندی تصادفی تولیدشده
۹: محاسبه مقدار خطای استاندارد برای مقادیر به‌دست‌آمده از شاخص به ازای تمام مدل‌های پیکربندی تصادفی تولیدشده
۱۰: محاسبه مقدار t برای شاخص
۱۱: ذخیره مقدار t در واحد متناظر آن در ماتریس مقادیر آزمون t
۱۲: **پایان حلقهٔ برای**
۱۳: **پایان حلقهٔ برای**
۱۴: **بازگردان** ماتریس مقادیر آزمون t

[44] Standard Error
[45] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ttest_1samp.html

۳.۲.۴.۳ مقایسه گراف‌های معنایی با استفاده از مقادیر آزمون آماری

پس از محاسبه مقادیر t به ازای هر شبکه و هر شاخص، برای بررسی اینکه هر شبکه به طرز معناداری با گراف‌های متناظر تصادفی خودش متفاوت هست یا خیر، مقادیر احتمال[۴۶] مربوطه محاسبه می‌شوند. با در نظر گرفتن یک گراف معنایی و هر یک از شاخص‌ها، فرض صفر و فرض جایگزین در این بخش به شرح زیر می‌باشند:

فرض صفر گراف معنایی G در شاخص X با مدل پیکربندی تصادفی خود تفاوت معناداری ندارد.

فرض جایگزین گراف معنایی G در شاخص X با مدل پیکربندی تصادفی خود به طرز معناداری متفاوت است.

به این صورت، به ازای تمامی شاخص‌ها، مقادیر احتمال با استفاده از مقادیر حاصل از آزمون آماری t برای هر گراف معنایی محاسبه می‌شود. با استفاده از این مقادیر می‌توان معنادار بودن تفاوت هر گراف با مدل‌های تصادفی متناظرش را مورد بررسی قرار داد.

میدانیم که مقادیر احتمال به تنهایی نمی‌توانند میزان تفاوت میان هر گراف و مدل‌های تصادفی آن را نشان دهند، و تنها گویای معنادار بودن آماری این تفاوت است. از آن‌جایی که هدف اصلی در این پژوهش طبقه‌بندی بازنمایی‌های معنایی غیرهم‌اندازه است، برای مقایسه نهایی میزان تفاوت هر گراف با مدل‌های تصادفی آن، نیاز به محاسبه مقادیر اندازه تاثیر[۴۷] است. این مقدار با استفاده از رابطه زیر محاسبه می‌شود:

$$\mathbf{d} = \frac{t}{\sqrt{n}} = \frac{\overline{X} - \mu}{\sigma} \qquad (۹.۳)$$

اندازه تاثیر در واقع نشان دهنده میزان تفاوت میان یک مشاهده و میانگین نمونه‌ها است. در این پژوهش مقدار یک شاخص برای یک گراف معنایی همان مشاهده ما است که با میانگین مقادیر شاخص برای ده مدل تصادفی گراف معنایی مقایسه می‌شود. با استفاده از رابطه بالا، ماتریس مربوط به مقادیر تاثیر ساخته می‌شود که به منظور طبقه‌بندی بازنمایی‌های معنایی از آن استفاده خواهیم کرد.

[46] Probability Value
[47] Effect Size

۴.۲.۴.۳ طبقه‌بندی بازنمایی‌های معنایی

پس از ساخت ماتریس مقادیر اندازه تاثیر، می‌توانیم به طبقه‌بندی بازنمایی‌های معنایی بر اساس ویژگی‌های ساختاری آن‌ها بپردازیم. ستون‌های این ماتریس نمایانگر شاخص‌های سراسری شبکه و سطرهای آن نمایانگر گراف‌های معنایی مورد مطالعه است. به منظور انجام طبقه‌بندی از روش خوشه‌بندی تجمعی ترتیبی[۴۸] استفاده می‌کنیم. این روش خوشه‌بندی، ابتدا برای هر گراف معنایی یک خوشه در نظر می‌گیرد و سپس با محاسبه شباهت میان مقادیر اندازه تاثیر برای هر گراف، گراف‌های معنایی مشابه را شناسایی می‌کند. حاصل این الگوریتم یک دندروگرام[۴۹] است که امکان مقایسه بازنمایی‌های معنایی را به دست می‌دهد.

۳.۴.۳ مطالعه گراف‌های معنایی در سطح میانی

همان‌طور که دیدیم، شاخص‌های مقایسه شبکه‌ها در مقیاس سراسری می‌توانند اطلاعات متنوعی درباره هر شبکه به دست دهند. با این وجود، جنبه‌های دیگری از ساختار و رفتار هر یک از بازنمایی‌های معنایی وجود دارد که نه در مقیاس سراسری، بلکه در مقیاس زیرشبکه‌های آنها نمود پیدا می‌کند. از طرفی، مطالعه ساختار و خصوصیات زیرشبکه‌های هر یک از گراف‌های معنایی، ممکن است در درک رفتار هر یک از بازنمایی‌ها راه‌گشا باشد. در ادامه، به توضیح نحوه مطالع جوامع در گراف‌های معنایی می‌پردازیم.

به طور کلی، یک جامعه در یک ساختار شبکه‌ای از گره‌ها را می‌توان به عنوان مجموعه‌ای از گره‌ها که با یک‌دیگر اتصالات بسیار و با سایر گره‌ها اتصالات کمی دارند تعریف کرد به بیان دیگر جامعه زیرشبکه‌ای است که در آن اتصالات زیاد بین موجودیت‌ها ساختاری متراکم درون شبکه ایجاد کرده است. معمولا، این نوع تراکم اتصالات در یک جامعه می‌تواند نشانگر شباهت موجودیت‌های حاضر در آن جامعه باشد.

الگوریتم مورد استفاده برای شناسایی جوامع در گراف‌های معنایی به منظور شناسایی جوامع در یک شبکه، علی‌الخصوص در سال‌های اخیر الگوریتم‌های متنوعی ارائه شده است که هر یک ممکن است شناسایی جوامع را با کیفیت متفاوتی انجام دهند. در این پژوهش، با توجه به تعداد زیاد گره‌ها و اتصالات در هر یک از شبکه‌های معنایی، از الگوریتم کلاسیک [۷۱] استفاده میکنیم چرا که این روش پیچیدگی محاسباتی و زمانی کمی دارد.

[48] Hierarchical Agglomerative Clustering
[49] Dendrogram

نحوه مقایسه جوامع در شبکه‌های معنایی به منظور مقایسه بازنمایی‌های معنایی از جهت احتمال تشکیل جوامع معنایی در آن‌ها نیاز است تا از شاخصی مرتبط استفاده کنیم. به همین علت شاخص پودمانگی[50] [۷۲]را برای هر گراف معنایی محاسبه می‌کنیم. شاخص پودمانگی درواقع نشان می‌دهد که تا چه اندازه امکان تقسیم یک گراف به پودمان‌های (جوامع و یا خوشه‌ها) کوچک‌تر وجود دارد. هرچه این امتیاز بالاتر باشد، نشان می‌دهد که گراف مورد نظر اتصالات زیاد درون-پودمانی و اتصالات محدود میان-پودمانی دارد. امتیاز پودمانگی (Q) با استفاده از رابطه زیر محاسبه می‌شود:

$$Q = \sum_{c=1}^{n} \left[\frac{L_c}{m} - \gamma \left(\frac{k_c}{2m} \right)^2 \right] \qquad (۱۰.۳)$$

در این رابطه، $\mathbf{m}$ تعداد کل یال‌های گراف، $\mathbf{L_c}$ تعداد اتصالات در جامعه $\mathbf{c}$، $\mathbf{k_c}$ مجموع درجات در جامعه $\mathbf{c}$ و γ مقدار دقت است که نسبت وزن اتصالات درون-پودمانی به اتصالات میان-پودمانی را تعیین می‌کند و معمولا برابر یک در نظر گرفته می‌شود. در فصل نتایج ضمن شناسایی جوامع در هر یک از شبکه‌های معنایی و مقایسه شاخص پودمانگی در آن‌ها، نمونه‌ای از جوامع به دست‌آمده نیز تصویر خواهد شد.

[50] Modularity Score

فصل ۴

نتایج

۱.۴ مقدمه

در فصل گذشته بازنمایی‌های معنایی که در این پژوهش مورد مطالعه قرار گرفته‌اند معرفی شدند و روش نگاشت هر بازنمایی معنایی به گراف متناظر آن نیز توضیح داده شد. سپس، شاخص‌های مبتنی بر علم شبکه که از آن‌ها استفاده می‌کنیم در دو سطح سراسری و میانی شرح داده شد. در نهایت، روش طبقه‌بندی گراف بازنمایی‌های معنایی با رویکرد آماری نیز برای خواننده توضیح داده شد.

در فصل جاری، ابتدا مقادیر به دست آمده برای هر یک از شاخص‌های سراسری برای شبکه‌های معنایی ارائه شده و الگوهای مشاهده شده گزارش شده است. سپس، نتایج آزمون آماری مقایسه گراف معنایی با مدل‌های تصادفی ارائه شده و با توجه به آن به خوشه‌بندی گراف‌های معنایی پرداخته‌ایم. در این قسمت، تفاوت خوشه‌بندی مبتنی بر آزمون آماری و خوشه‌بندی مبتنی بر مقادیر شاخص‌های سراسری شبکه بررسی می‌شود. در این میان، نحوه توزیع درجات در شبکه‌های معنایی نیز تصویر شده و رئوس تاثیرگذار هر گراف معنایی شناسایی شده است. با توجه به این نتایج، میزان همپوشانی رئوس تاثیرگذار گراف‌های معنایی با کلمات پرتکرار زبان انگلیسی بررسی شده است. در آخر، ساختار میانی گراف‌های معنایی با مقایسه امتیاز پودمانگی آن‌ها مورد بررسی قرار گرفته است.

۲.۴ گراف‌های حاصل از نگاشت بازنمایی‌های معنایی

مطابق مراحل ذکر شده در فصل قبل گراف معنایی حاصل از هر بازنمایی تشکیل شده است. لازم به ذکر است که بازنمایی‌های معنایی که از فضای برداری گرفته شده‌اند هرکدام به پنج گراف مستقل نگاشت شده‌اند. این تفکیک درواقع بر اساس پنج آستانه تشابه از پیش تعیین شده در بازه‌ی ۰٫۵ تا ۰٫۷ صورت گرفته است. این در حالی است که بازنمایی‌های معنایی مبتنی بر پایگاه داده واژگانی نگاشتی یک به یک دارند. با این توصیف، در مجموع بیست و سه گراف معنایی حاصل می‌شود که در ادامه به نتایج حاصل از مقایسه آن‌ها می‌پردازیم.

۳.۴ ویژگی‌های سراسری شبکه‌های معنایی

اولین دسته از شاخص‌های شبکه مورد استفاده در چهارچوب پیشنهادی، شاخص‌های مقیاس سراسری هستند. شاخص‌های پایه گزارش شده برای هر گراف عبارت هستند از تعداد رئوس، تعداد یال و تعداد اجزای متصل در هر گراف. در قدم بعدی به گزارش مقادیر به دست آمده برای شاخص‌های میانگین درجات، مقدار بزرگترین درجه، قطر، متوسط فاصله، ضریب خوشگی سراسری، متوسط ضریب خوشگی سراسری و ضریب همسان‌گرایی درجه‌ای می‌پردازیم.

۱.۳.۴ شاخص‌های سراسری گراف‌های مبتنی بر دانش انسانی

به منظور مقایسه شاخص‌های سراسری میان سه گراف وردنت، فرهنگ موبی و کانسپت‌نت مقادیر مربوطه در جدول ۱.۴ ارائه شده است.

جدول ۱.۴: مشخصات ساختاری سراسری بازنمایی‌های معنایی مبتنی بر دانش انسانی

ویژگی	وردنت	موبی	کانسپت‌نت
تعداد اجزای متصل	۲۹۵۰۰	۱	۳۲۵۶۱
تعداد گره	۳۲۶۱۱	۱۰۷۹۸۰	۲۹۶۶۹۶
تعداد یال	۱۱۹۴۶۳	۱۷۹۹۹۳۳	۳۷۶۳۳۰
بزرگترین درجه	۱۵۲	۱۴۶۸	۳۷۳۸
تراکم یال	۲×۱۰^{-۴}	۳×۱۰^{-۴}	$۸/۵۵ \times ۱۰^{-۶}$
میانگین درجات	۷/۳۲	۳۳/۳۳	۲/۵۳
قطر	۲۳	۹	۲۷
متوسط فاصله	۶/۸۹	۳/۸۶	۷/۲۳
ضریب خوشگی سراسری	۰/۳۶	۰/۱۹	۰/۰۰۲۸
متوسط ضریب خوشگی محلی	۰/۶۲	۰/۶۵	۰/۰۲
ضریب همسانگرایی درجه‌ای	۰/۲۶	۰/۰۳	−۰/۰۲

از میان این سه گراف معنایی حاصل، تنها موبی گرافی متصل است و دو گراف دیگر از چند زیرگراف غیرمتصل تشکیل شده‌اند. لذا در این پژوهش به منظور ایجاد امکان مقایسه، بزرگترین جزء وردنت و بزرگترین جزء کانسپت‌نت در نظر گرفته شده است.

اولین تفاوت قابل ملاحظه در این جدول، تفاوت میان مقادیر به دست آمده برای شاخص تراکم یال در این شبکه‌ها است. این شاخص کمترین میزان خود را در کانسپت‌نت و بیشترین آن را در موبی نشان می‌دهد. این مشاهده حاکی از میزان دربرگیری روابط معنایی بالقوه در هر یک از این شبکه‌های معنایی است و با توجه به آنچه در فصول قبلی درباره نحوه ساخت هر یک از این بازنمایی‌های معنایی گفته شد، می‌توان این مشاهده را توضیح داد.

می‌دانیم که در هر دو بازنمایی وردنت و موبی عمدتا وجود و یا عدم ارتباط میان دو کلمه بر پایه رابطه هم‌معنایی تعریف می‌شود. هرچند در موبی، رابطه هم‌معنایی تعریف وسیع‌تر و دربرگیرنده‌تری دارد و به همین علت نیز این بازنمایی روابط معنایی بیشتری را در مقایسه با وردنت دربرمی‌گیرد و گراف حاصل نیز تراکم یال بیشتری دارد. از طرفی، علی رغم اینکه در بازنمایی کانسپت‌نت روابط معنایی بسیار متنوعی در نظر گرفته شده‌اند، گراف حاصل به مراتب تنک‌تر از دو گراف ذکر شده است. می‌دانیم که در کانسپت‌نت بالغ بر سی نوع[1] رابطه معنایی در نظر

[1] https://github.com/commonsense/conceptnet5/wiki/Relations

گرفته شده است که رابطه هم‌معنایی تنها یکی از آن‌ها است. با این وجود، ممکن است به علت در نظر نگرفتن بخش غیرانگلیسی زبان کانسپت‌نت، گراف مورد مطالعه در این پژوهش تراکم خود را از دست داده باشد.

این مسئله در مقایسه متوسط درجات این سه گراف معنایی نیز مشهود است.

همانطور که در جدول مشاهده می‌شود، تفاوت این سه گراف در قطر و متوسط فاصله آنها نیز قابل توجه است. در هر دو این شاخص‌ها کاسپت‌نت و موبی به ترتیب بزرگ‌ترین و کوچک‌ترین این مقادیر را به خود اختصاص داده‌اند.

از طرف دیگر، با نظر به مقادیر به دست آمده برای شاخص‌های ضرایب خوشگی، می‌توان نتیجه گرفت که احتمال تشکیل سه‌تایی بسته در کانسپت‌نت به میزان قابل توجهی پایین‌تر از دو گراف همتای آن است. دو گراف دیگر، وردنت و موبی ، ضریب خوشگی سراسری یکسانی دارند اما موبی ضریب خوشگی محلی پایین‌تری دارد. [۱۰] این مشاهده را به این شکل توضیح می‌دهند که رئوس بیشتری در موبی دارای درجه بالا هستند و در گراف‌های واقعی معمولا رئوس با درجه بالا ضریب خوشگی پایین‌تری دارند. از طرفی، در تعریف ضریب خوشگی سراسری، رئوس با درجه بالاتر، سهم بیشتری دارند که این موضوع در مقدار حاصل برای موبی نیز خود را نشان می‌دهد.

با در نظر گرفتن مقادیر گزارش شده، [۱۰] نتیجه گرفته‌اند که دو شبکه وردنت و موبی ویژگی‌هایی شبیه به شبکه‌های واقعی مطالعه شده پیشین دارند [۷۳]. به علاوه، با توجه به کوچک بودن متوسط فاصله و قطر و بزرگ بودن ضریب خوشگی برای این دو شبکه، نتیجه گرفته‌اند که این شبکه‌ها جهان‌کوچک[۲] هستند.

شاخص سراسری نهایی که می‌تواند تمایز این شبکه‌ها را بیشتر نمایان کند، ضریب همسان‌گرایی درجه‌ای است[۳]. مقدار این شاخص برای وردنت نشان می‌دهد که از میان سه گراف مورد مطالعه، رئوس این شبکه بیشترین میزان تمایل را به اتصال با رئوس هم‌درجه خود دارند. اما این شاخص، برای موبی بسیار کمتر است که نشان می‌دهد رئوس آن تمایل چندانی به اتصال به رئوس هم‌درجه خود را ندارند. در این میان کانسپت‌نت ضریب همسان‌گرایی درجه‌ای منفی دارد که نشان‌دهنده وجود تمایل هرچند اندک میان رئوس این شبکه برای متصل شدن به رئوسی با درجه متفاوت از درجه خودشان است.

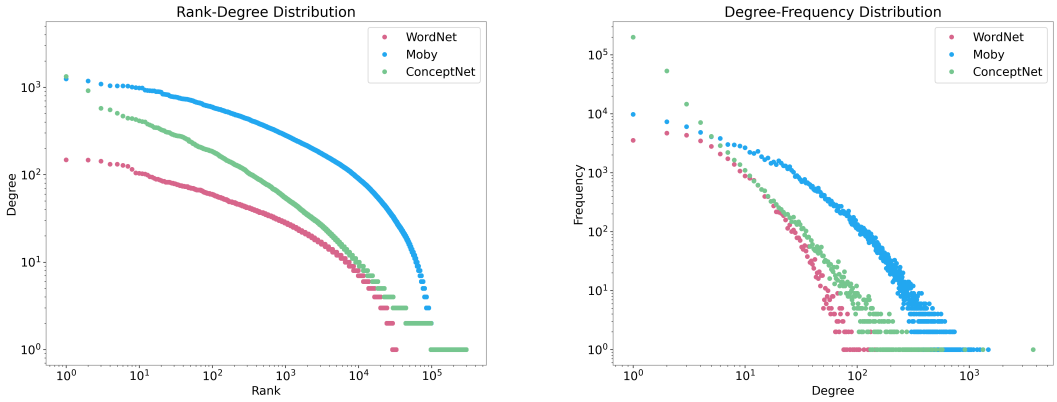
با در نظر گرفتن مقادیر حاصل از شاخص‌های سراسری در کانسپت‌نت، با توجه به اینکه منابع مورد استفاده برای ساخت این بازنمایی معنایی همانند منابع وردنت و موبی متکی بر دانش انسانی هستند، انتظار می‌رفت که

[2] Small-world

[۳] توضیح کامل این شاخص در فصل سوم ارائه شده است.

الگوهای مشابهی در نتایج شاخص‌های سراسری این گراف و دو گراف دیگر مشاهده کنیم. در حالی که تقریبا در تمامی شاخص‌ها، تفاوت‌های عمده‌ای میان نتایج کانسپت‌نت و دو گراف دیگر وجود دارد. همانطور که اشاره شد، مقادیر نسبتا بزرگ‌تری برای قطر و متوسط فاصله در این گراف به دست آمده است. از طرفی، در صورت انتخاب یک راس تصادفی در این شبکه، احتمال اینکه همسایگان آن راس به یکدیگر متصل باشند بسیار پایین و نزدیک به صفر است. این مشاهدات، احتمال جهان‌کوچک بودن این شبکه را به شدت کاهش می‌دهد. از طرفی دیگر، کوچکی مقدار ضریب همسان‌گرایی درجه‌ای در این شبکه، علی‌الخصوص با در نظر گرفتن کوچکی مقدار ضرایب خوشگی آن، می‌توان گفت که اتصالات در این شبکه رفتاری به نسبت تصادفی‌تری دارند و الگویی در آن مشاهده نشد.

۱.۱.۳.۴ توزیع درجات در گراف‌های معنایی مبتنی بر دانش انسانی



(آ) نمودار درجه–فراوانی (ب) نمودار رتبه–درجه

شکل ۱.۴: نمودار توزیع درجات در شبکه‌های معنایی مبتنی بر پایگاه داده واژگان

در تصویر ۱.۴ نمودار توزیع درجات هر سه شبکه معنایی مورد بحث در این قسمت در دو قالب توزیع درجه بر حسب فراوانی و همچنین توزیع رتبه هر راس بر حسب درجه آن ارائه شده است. آنجه از ۱.۴.آ برمی‌آید این است که در تمامی این شبکه‌ها، توزیع درجات میان رئوس نه تنها متقارن نیست بلکه انحراف مثبت دارد. این مشاهده نشان می‌دهد که رئوس با درجات پایین فراوانی بیشتری نسبت به رئوس با درجات بالاتر دارند. از آنجایی

که این الگو در شبکه‌های واقعی امری رایج است، مشاهده آن در میان این سه شبکه معنایی انسانی نیز مورد انتظار بود.

از طرفی، پس از بررسی توزیع درجات این سه گراف، مشاهده شد که در دو گراف موبی و کانسپت‌نت، با در نظر گرفتن بخشی از محدوده داده [۴]، می‌توانند توزیع توانی در نظر گرفته شوند. این موضوع درباره وردنت صادق نیست، و تفاوت معناداری میان متناسب بودن توزیع‌های رایج برای توزیع درجات این گراف وجود ندارد.

در ادامه به بررسی و مقایسه نتایج حاصل از به‌کارگیری معیارهای سراسری شبکه در گراف‌های حاصل از بازنمایی‌های معنایی برداری می‌پردازیم.

۲.۳.۴ مقایسه شاخص‌های سراسری میان گراف‌های معنایی مبتنی بر مدل‌سازی توزیعی معنا

در قسمت قبل به مقایسه نتایج شاخص‌های سراسری میان شبکه‌های معنایی انسانی پرداختیم. در این بخش، ضمن بررسی این شاخص‌ها در شبکه‌های معنایی مبتنی بر یادگیری ماشینی، به مقایسه این نتایج، با قسمت پیشین پرداخته شده است.

نتایج شاخص‌های سراسری برای هر یک از بازنمایی‌هایی معنایی ازپیش‌آموزش‌دیده در جدول‌های زیر ارائه شده‌اند.

[4] Truncated Power Law

جدول ۲.۴: مشخصات ساختاری بازنمایی معنایی ازپیش‌آموزش‌دیده Word2Vec

ویژگی	۰/۵	۰/۵۵	۰/۶	۰/۶۵	۰/۷
مشخصات تمامی اجزاء					
تعداد گره	۵۸۱۸۶	۵۰۵۷۶	۳۹۳۶۳	۲۶۵۰۹	۱۴۶۹۴
تعداد یال	۲۰۳۳۲۹۷	۵۰۱۰۸۵	۲۷۲۴۵۸	۷۸۲۲۷	۱۹۵۵۷
میانگین درجات	۶۹/۸۸	۳۱/۶۷	۱۳/۸۴	۵/۹۰	۲/۶۶
مشخصات بزرگ‌ترین جزء متصل					
تعداد گره	۵۷۱۰۲	۴۶۷۱۷	۲۹۳۷۴	۱۱۳۶۳	۱۷۳۹
تعداد یال	۲۰۳۲۵۳۰	۷۹۸۳۶۳	۲۶۳۷۳۱	۶۲۸۳۴	۵۸۸۵
بزرگترین درجه	۱۸۶۶	۱۱۱۷	۶۱۳	۳۲۸	۱۵۹
میانگین درجات	۷۱/۱۸	۳۴/۱۷	۱۷/۹۵	۱۱/۰۵	۶/۷۶
تراکم یال	۰/۰۰۱	۰/۰۰۰۷	۰/۰۰۰۶	۰/۰۰۰۹	۰/۰۰۳
قطر	۲۱	۲۷	۴۵	۶۷	۲۳
متوسط فاصله	۵/۷۴	۷/۸۶	۱۱/۳۸	۱۸/۱۷	۸/۶۲
ضریب خوشگی سراسری	۰/۴۳	۰/۴۳	۰/۴۱	۰/۳۹	۰/۲۸
متوسط ضریب خوشگی محلی	۰/۳۷	۰/۳۶	۰/۳۶	۰/۳۶	۰/۳۲
ضریب همسانگرایی درجه‌ای	۰/۴۳	۰/۴۱	۰/۴	۰/۳۷	۰/۱۱

می‌دانیم که هر بازنمایی معنایی ازپیش‌آموزش‌دیده با توجه به آستانه‌های تشابه در نظر گرفته شده به پنج گراف مستقل نگاشت شده است. به همین علت، برای مقایسه نتایج شاخص‌ها، مقادیر به‌دست‌آمده میان پنج گراف حاصل از هر بازنمایی مقایسه می‌شود. از طرفی، مقایسه این مقادیر میان بازنمایی‌های ازپیش‌آموزش‌دیده مختلف و بازنمایی‌های انسانی نیز مقایسه و بررسی می‌شوند.

اولین مورد جالب توجه در نتایج به دست آمده از شاخص‌های سراسری، مقادیر مربوط به شاخص میانگین درجات می‌باشد. در جدول نخست که مقادیر شاخص‌ها را برای گراف‌های Word2Vec نشان می‌دهد، می‌توان دید که با هر بار افزایش آستانه تشابه، مقدار میانگین درجات تقریبا نصف می‌شود. می‌توان گفت که رابطه میانگین درجات با آستانه تشابه رابطه‌ای توانی[5]است. نکته قابل توجه این است که این الگو در رابطه میان میانگین درجات و آستانه تشابه، تقریبا در تمامی گراف‌های معنایی دیگر مبتنی بر یادگیری ماشین نیز قابل مشاهده است و تنها گرافی که این خاصیت را نشان نمی‌دهد، ConceptNetNumberbatch می‌باشد.

[5]Power Law

جدول ۳.۴: مشخصات ساختاری بازنمایی معنایی ازپیش‌آموزش‌دیده BERT2Static

ویژگی	۰/۵	۰/۵۵	۰/۶	۰/۶۵	۰/۷
مشخصات تمامی اجزاء					
تعداد گره	۴۸۷۲۳	۴۴۳۵۰	۳۷۵۹۳	۲۹۰۷۷	۱۹۶۵۶
تعداد یال	۲۹۲۰۴۲۱	۱۳۵۱۱۹۷	۵۹۳۱۷۲	۲۴۳۰۶۱	۸۷۷۷۹
میانگین درجات	۱۱۹/۸۷	۶۰/۹۳	۳۱/۵۵	۱۶/۷۱	۸/۹۳
مشخصات بزرگ‌ترین جزء متصل					
تعداد گره	۴۷۲۵۷	۴۰۷۰۹	۲۹۸۶۱	۱۷۸۱۶	۸۵۱۷
تعداد یال	۲۹۱۹۳۶۸	۱۳۴۸۱۵۴	۵۸۵۹۴۴	۲۲۹۵۲۶	۲۲۹۵۲۶
میانگین درجات	۱۲۳/۵۵	۶۶/۲۳	۳۹/۲۴	۲۵/۷۶	۱۷/۸۱
تراکم یال	۰/۰۰۲	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱	۰/۰۰۲
بزرگترین درجه	۱۹۸۸	۱۱۳۷	۷۰۰	۳۸۱	۲۲۳
قطر	۲۵	۳۰	۳۳	۳۹	۳۶
متوسط فاصله	۴/۸۸	۶/۰۹	۷/۵۹	۸/۸۵	۱۰/۵۷
ضریب خوشگی سراسری	۰/۴۳	۰/۴۵	۰/۴۹	۰/۵۱	۰/۵۱
متوسط ضریب خوشگی محلی	۰/۴۰	۰/۴۰	۰/۴۰	۰/۴۰	۰/۴۱
ضریب همسانگرایی درجه‌ای	۰/۴۲	۰/۵	۰/۵۵	۰/۵۷	۰/۵۶

با توجه به مقادیر به‌دست‌آمده شاخص تراکم یال برای هر یک از گراف‌ها، می‌بینیم که VisWord2Vec و کانسپت‌نت به ترتیب متراکم‌ترین و تنک‌ترین گراف‌ها هستند. با توجه به اینکه تراکم یال‌ها می‌تواند ناشی از در نظر گرفتن تعداد بیشتری از روابط بالقوه در یک گراف باشد، می‌توان گفت VisWord2Vec روابط معنایی زیادی را نسبت به سایر بازنمایی‌ها در خود لحاظ می‌کند. از طرفی، می‌دانیم که در مرحله آموزش VisWord2Vec علاوه بر داده متنی، از داده تصویری هم استفاده می‌شود.این نوع از آموزش به منظور در نظر گرفتن روابط معنایی بیشتر که ممکن است دستیابی به آن‌ها از طریق داده متنی به تنهایی میسر نباشد انجام شده است. با در نظر گرفتن این موضوع، بیشتر بودن تعداد متوسط ارتباطات در گراف حاصل از VisWord2Vec مورد انتظار است.

به طور کلی، گراف‌های حاصل از VisWord2Vec بیشترین میزان تراکم و گراف‌های کانسپت‌نت، وردنت، موبی و سه تا از گراف‌های ConceptnetNumberbatch کمترین میزان تراکم یال را دارند. نقطه اشتراک تمامی این گراف‌ها این است که منبع آن‌ها برای بازنمایی معنایی تماما و یا تا قسمتی مبتنی بر پایگاه داده واژگان است. می‌توان

این‌طور توضیح داد که پایگاه دادگان واژگانی که عموما توسط نیروی انسانی ساخته می‌شود به علت محدودیت‌های روش ساخت، ممکن است روابط معنایی زیادی را در بازنمایی نهایی لحاظ نکنند.

در نهایت، مشاهده روند تغییر مقادیر تراکم یال با افزایش آستانه تشابه نشان می‌دهد که در گراف‌های حاصل از دو بازنمایی VisWord2Vec و ConceptnetNumberbatch با هر بار کاهش آستانه تشابه، مقدار مربوط به تراکم یال کاهش می‌یابد. این در حالی است که در گراف‌های حاصل از Word2Vec و BERT2Static، روند نزولی تنها تا آستانه تشابه ۰/۷ ادامه دارد اما در این مقدار مجددا تراکم یال به مقدار اولیه خود و یا حتی بالاتر می‌رسد. به نظر می‌رسد این دو بازنمایی در اکثر شاخص‌ها برای آستانه تشابه ۰/۷ رفتاری متفاوت از سایر گراف‌ها نشان می‌دهند که زمان ارائه نتایج شاخص‌های دیگر به این موضوع خواهیم پرداخت.

حال به بررسی نتایج معیارهای مبتنی بر فاصله می‌پردازیم. اول از همه، با نظر به مقادیر مربوط به شاخص قطر، مشخص است که این شاخص در گراف موبی و گراف‌های VisWord2Vec کوچک‌ترین مقادیر خود را دارد. از طرفی، گراف‌های حاصل از دو آستانه تشابه ۰/۵۵ و ۰/۶ بازنمایی Word2Vec بیشترین مقدار قطر را دارند. گفتنی است که با هر بار افزایش مقدار آستانه تشابه، اندازه قطر در بازنمایی‌های VisWord2Vec و ConceptnetNumberbatch افزایش می‌یابد. هر چند در دو بازنمایی Word2Vec و BERT2Static به رغم افزایش کلی در هر بار افزایش تشابه، باز هم در آستانه ۰/۷ روند کلی شکسته شده و شاهد کاهش دوباره قطر هستیم. با توجه به تغییرات قطر در جدول اول، مشاهده می‌کنیم که با هر مرحله افزایش آستانه تشابه، قطر گراف Word2Vec افزایش چشم‌گیری دارد. این موضوع درباره متوسط فاصله در این گراف نیز صادق است.

جدول ۴.۴: مشخصات ساختاری بازنمایی معنایی ازپیش‌آموزش‌دیده VisWord2Vec

ویژگی	۰/۵	۰/۵۵	۰/۶	۰/۶۵	۰/۷
مشخصات تمامی اجزاء					
تعداد گره	۱۱۱۵۶	۱۰۸۹۶	۱۰۴۰۲	۹۵۷۶	۸۱۰۲
تعداد یال	۷۵۴۴۷۹۱	۴۴۶۷۱۵۶	۲۳۶۷۸۳۷	۱۱۰۰۶۷۶	۴۳۰۴۲۲
میانگین درجات	۱۳۵۲/۵۹	۸۱۹/۹۶	۴۵۵/۲۶	۲۲۹/۸۸	۱۰۶/۲۵
مشخصات بزرگ‌ترین جزء متصل					
تعداد گره	۱۱۱۵۶	۱۰۸۷۰	۱۰۳۳۷	۹۴۵۱	۷۸۸۰
تعداد یال	۷۵۴۴۷۹۱	۴۴۶۷۱۳۷	۲۳۶۷۷۹۶	۱۱۰۰۶۰۲	۴۳۰۲۸۹
میانگین درجات	۱۳۵۲/۵۹	۸۲۱/۹۲	۴۵۵/۲۶	۲۳۲/۹۰	۱۰۹/۲۱
تراکم یال	۰/۱۲	۰/۰۷	۰/۰۴	۰/۰۲	۰/۰۱
بزرگترین درجه	۶۰۳۵	۴۹۸۹	۳۸۳۸	۲۶۸۹	۱۶۰۳
قطر	۹	۱۰	۱۲	۱۳	۱۳
متوسط فاصله	۲/۱۶	۲/۳۹	۲/۶۳	۳/۰۷	۳/۵۷
ضریب خوشگی سراسری	۰/۵۵	۰/۵۲	۰/۴۹	۰/۴۸	۰/۴۶
متوسط ضریب خوشگی محلی	۰/۶۲	۰/۵۹	۰/۵۶	۰/۵۳	۰/۴۹
ضریب همسانگرایی درجه‌ای	۰/۰۷	۰/۱	۰/۱۴	۰/۱۸	۰/۱۹

از دیگر شاخص‌های مبتنی بر فاصله، متوسط فاصله میان رئوس است. مقادیر موجود در جدول ۴.۴ نشان می‌دهد که متوسط فاصله در گراف‌های مربوط به VisWord2Vec کمترین مقدار را دارد. بعد از این پنج گراف، گراف موبی نیز متوسط فاصله کوچکی دارد. مجددا گرافی متعلق به Word2Vec در آستانه تشابه ۰/۶۵، بزرگ‌ترین مقدار متوسط فاصله را دارد.

از طرفی در سه بازنمایی BERT2Static، VisWord2Vec و ConceptnetNumberbatch می‌توان دید که با هر بار افزایش آستانه تشابه، متوسط فاصله نیز افزایش می‌یابد. هر چند، مقدار متوسط فاصله در بازنمایی Word2Vec به رغم روند کلی افزایشی، در آستانه ۰/۷ مجددا کاهش می‌یابد.

به طور کلی می‌توان گفت که با توجه به مقادیر به دست آمده، در بازنمایی VisWord2Vec فاصله میان مفاهیم به طور متوسط از تمامی بازنمایی‌های دیگر به طرز چشم‌گیری کمتر است. این مشاهده در کنار مقادیر گزارش شده میزان تراکم یال نشان می‌دهد که این گراف با در نظر گرفتن اطلاعات معنایی به دست آمده از دادگان تصویری، موفق شده است روابط معنایی بیشتری را در بازنمایی نهایی لحاظ کند و به همین علت، جابه‌جایی از

یک کلمه به کلمه‌ای دیگر در این گراف راحت‌تر صورت می‌گیرد.

جدول ۵.۴: مشخصات ساختاری بازنمایی معنایی ازپیش‌آموزش‌دیده Conceptnet Numberbatch

ویژگی	۰/۵	۰/۵۵	۰/۶	۰/۶۵	۰/۷
مشخصات تمامی اجزاء					
تعداد گره	۸۷۷۲۸	۸۶۸۲۰	۸۴۹۱۴	۸۱۶۸۳	۷۶۳۶۲
تعداد یال	۳۳۱۳۹۲۲	۱۸۶۹۷۴۷	۱۰۴۵۵۱۳	۵۸۹۲۲۰	۳۳۷۲۰۶
میانگین درجات	۷۵/۵۴	۴۳۰/۷	۲۴/۶۲	۱۴/۴۲	۸/۸۳
مشخصات بزرگ‌ترین جزء متصل					
تعداد گره	۸۷۴۸۹	۸۶۰۸۶	۸۲۶۴۰	۷۴۷۸۲	۵۸۸۲۷
تعداد یال	۳۳۱۳۷۶۱	۱۸۶۹۱۶۸	۱۰۴۳۵۸۰	۵۸۲۱۳۹	۳۱۷۵۸۴
میانگین درجات	۷۵/۷۵	۴۳/۴۲	۲۵/۲۵	۱۵/۵۶	۱۰/۷۹
تراکم یال	۰/۰۰۰۸	۰/۰۰۰۵	۰/۰۰۰۳	۰/۰۰۰۲	۰/۰۰۰۱
بزرگترین درجه	۱۷۲۴	۱۱۵۱	۷۴۲	۴۹۲	۳۳۱
قطر	۱۴	۱۷	۲۲	۳۴	۴۷
متوسط فاصله	۴/۹۳	۶/۰۱	۴/۲۸	۹/۸۹	۱۳/۸۸
ضریب خوشگی سراسری	۰/۴۲	۰/۴۴	۰/۴۶	۰/۵۱	۰/۶
متوسط ضریب خوشگی محلی	۰/۴۷	۰/۵	۰/۵۱	۰/۵۱	۰/۵۱
ضریب همسانگرایی درجه‌ای	۰/۵۵	۰/۵۵	۰/۵۷	۰/۶۲	۰/۷۳

دسته بعدی شاخص‌های سراسری که در این پژوهش مورد استفاده قرار گرفته‌اند، ضرایب خوشه‌بندی هستند. شاخص خوشه‌بندی سراسری و متوسط شاخص خوشه‌بندی محلی نشان می‌دهند تا چه اندازه احتمال تشکیل خوشه[۶] در یک گراف وجود دارد. کوچک‌ترین میزان خوشه‌بندی سراسری و متوسط خوشه‌بندی محلی هر دو متعلق به کانسپت‌نت است. از طرفی، با اینکه در مقادیر به دست آمده برای ضریب خوشه‌بندی سراسری، موبی و وردنت نیز مقادیر بسیار کوچکی دارند، اما در متوسط ضریب خوشه‌بندی محلی، این دو گراف بالاترین مقادیر را به خود اختصاص داده‌اند. بنابراین می‌توان این‌طور نتیجه گرفت که گراف‌های معنایی مبتنی بر پایگاه داده واژگان، ضریب خوشه‌بندی سراسری بسیار کمی دارند اما این امر در مورد ضریب خوشه‌بندی محلی لزوما صادق نیست.

از طرفی، در دو بازنمایی Word2Vec و VisWord2Vec با افزایش آستانه تشابه، میزان ضریب خوشگی سراسری کاهش می‌یابد. اما در بازنمایی Conceptnet Numberbatch با هر بار افزایش آستانه تشابه، مقدار

[6] Cluster

ضریب خوشگی سراسری افزایش می‌یابد که این روند به طور کلی در BERT2Static هم مشاهده می‌شود.

شاخص نهایی که در این بخش به آن پرداخته شده است، شاخص همسان‌گرایی درجه‌ای می‌باشد. اولین مشاهده قابل بحث در این قسمت این است که تنها گرافی که مقدار همسان‌گرایی منفی دارد، کانسپت‌نت است. می‌دانیم که این شاخص، مقداری در بازه (۱،۱−) دارد و مقادیر منفی نشان‌دهنده تمایل رئوس به متصل شدن به رئوسی با درجه کاملا متفاوت از درجه خودشان است. با این توضیح، تنها گرافی که در آن گره‌های با درجات بزرگ‌تر تمایل بیشتری به اتصال با گره‌های با درجات پایین‌تر دارد، همین گراف کانسپت‌نت است. ممکن است در چنین گرافی، به دلیل وجود اتصالات میان رئوس با بزرگی درجه مخالف، انتقال اطلاعات راحت‌تر صورت بگیرد. از طرفی، از آنجایی که هر چه مقدار مطلق این ضریب به صفر نزدیک‌تر باشد، می‌توان گفت که در اتصال رئوس الگوی خاصی وجود ندارد و اغلب اتصالات، رفتاری تصادفی دارند. به همین علت دو گراف موبی و کانسپت‌نت که مقدار مطلق ضریب همسان‌گرایی کوچکی دارند به طور کلی ممکن است ساختار اتصالات آنها تصادفی‌تر از سایر بازنمایی‌ها باشد.

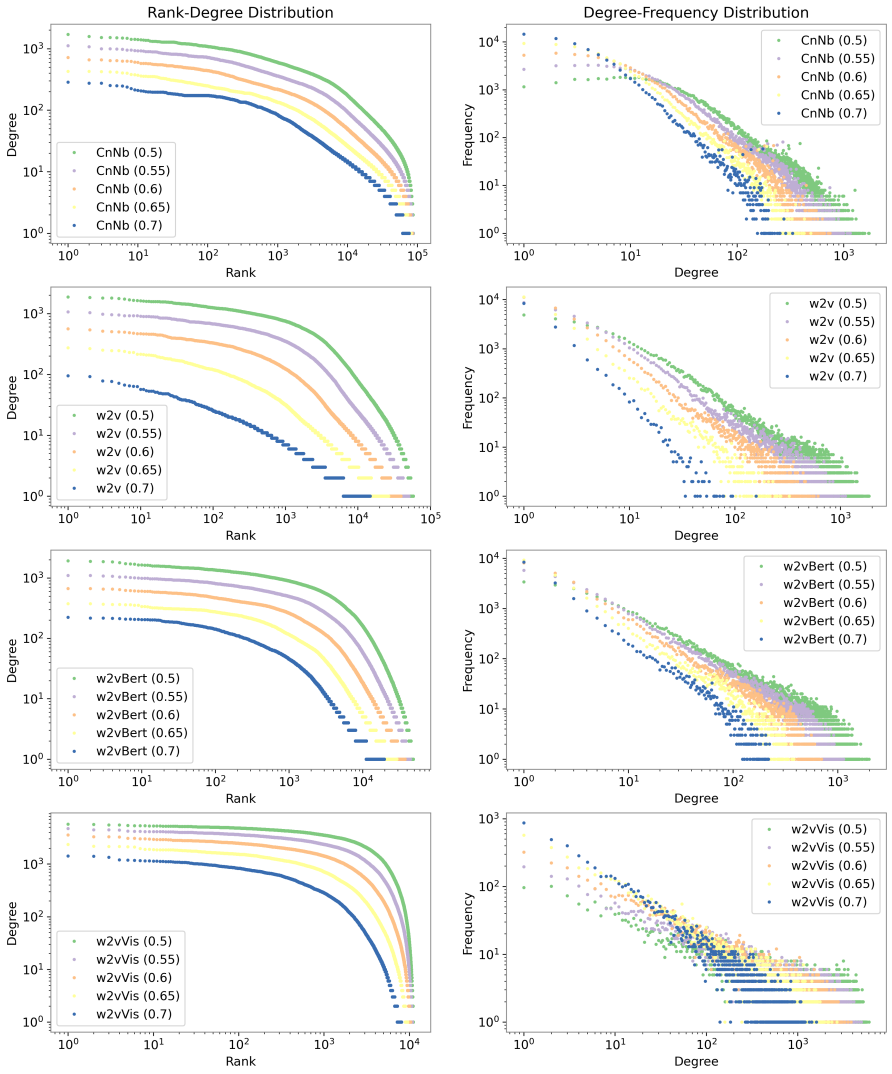
به منظور مقایسه روند تغییر این شاخص در ارتباط با آستانه تشابه می‌توان گفت که در بازنمایی Word2Vec با افزایش مقدار آستانه، مقدار این شاخص کاهش می‌یابد. این در حالی است که با افزایش آستانه تشابه در دو بازنمایی VisWord2Vec و Conceptnet Numberbatchمقدار این شاخص کاهش می‌یابد.

۱.۲.۳.۴ توزیع درجات در گراف‌های معنایی مبتنی بر مدل‌سازی توزیعی معنا

توزیع درجات گراف‌های معنایی Word2Vec، BERT2Static، VisWord2Vec و Conceptnet Numberbatch در تصویر ۴.۴ قابل مشاهده است. مشابه آنچه در تصویر ۱.۴ دیدیم، باز هم انحراف[7] مثبت در توزیع آماری این شبکه‌های معنایی قابل مشاهده بوده و این موضوع میان همه گراف‌ها مشترک است. هرچند توزیع درجات در این شبکه‌ها یکسان نیست. برای روشن شدن این موضوع، به بررسی این توزیع‌ها می‌پردازیم. در همه گراف‌های حاصل از دو بازنمایی معنایی Word2Vec و Conceptnet Numberbatch، در محدوده تقطیع شده توزیع درجات آنها، توزیع توانی را می‌توان مشاهده کرد. این در حالی است که گراف‌های بازنمایی BERT2Static بیشتر نمایی کشیده دارند و تنها در دو آستانه تشابه ۰٫۶۵ و ۰٫۷ توزیع توانی از خود نشان می‌دهند. گراف‌های حاصل از VisWord2Vec نیز همگی یا توزیع نمایی و یا توزیع نمایی کشیده دارند. لازم به ذکر است که این نتایج

[7] Skewness

بر حسب آزمون آماری و با مقدار احتمال[8] کم‌تر از ۰/۱ به دست آمده‌اند. همچنین، برای هیچ‌یک از گراف‌های معنایی مورد مطالعه در این پژوهش، توزیع توانی برای همه رئوس مشاهده نشد.



شکل ۲.۴: نمودار توزیع درجات در شبکه‌های معنایی حاصل از مدل‌های توزیعی معنا

[8]Probability Value

۳.۳.۴ طبقه‌بندی گراف‌های معنایی با استفاده از مقادیر شاخص‌های سراسری

با توجه به هدف اصلی این پژوهش که انجام مقایسه میان بازنمایی‌های معنایی و آشکار ساختن شباهت‌های آن‌هاست، در ادامه یک طبقه‌بندی از گراف‌های معنایی که با استفاده از مقادیر به دست آمده برای هر شاخص انجام شده، ارائه شده است. در این بخش از یک روش خوشه‌بندی سلسله‌مراتبی[9] استفاده شده است که در بخش قبلی به توضیح آن پرداخته‌ایم. در تصویر ۳.۴ نمایی از این خوشه‌بندی ارائه شده است. مقادیر موجود در ماتریس با روش کمینه-بیشینه[10] نرمال شده‌اند تا بازنمایی طبقه‌بندی نهایی قابل فهم باشد. خلاصه تمامی تحلیل‌هایی که در این بخش درباره مقایسه گراف‌ها گفته شد را می‌توان در این تصویر مشاهده کرد.

در اینجا قصد داریم به بررسی نتایج حاصل از خوشه‌بندی معرفی شده بپردازیم. در بالای تصویر، خوشه‌بندی مربوط به خود شاخص‌های سراسری را می‌توان مشاهده کرد و با توجه به وجود یک همبستگی کلی میان نتایج ضرایب خوشه‌بندی محلی و سراسری، می‌توان دید که این دو ضریب در یک گروه قرار گرفته‌اند. حال، با در نظر گرفتن خوشه‌بندی ارائه شده در سمت راست تصویر، می‌توان گروه‌بندی گراف‌های معنایی با یک‌دیگر را بررسی کرد. اولین شباهت بارز در این خوشه‌بندی، شباهت میان دو گراف موبی و وردنت است. با در نظر گرفتن تمامی مقادیر گزارش شده برای این دو گراف، به نظر می‌رسد که شباهت قابل توجهی به یکدیگر دارند.

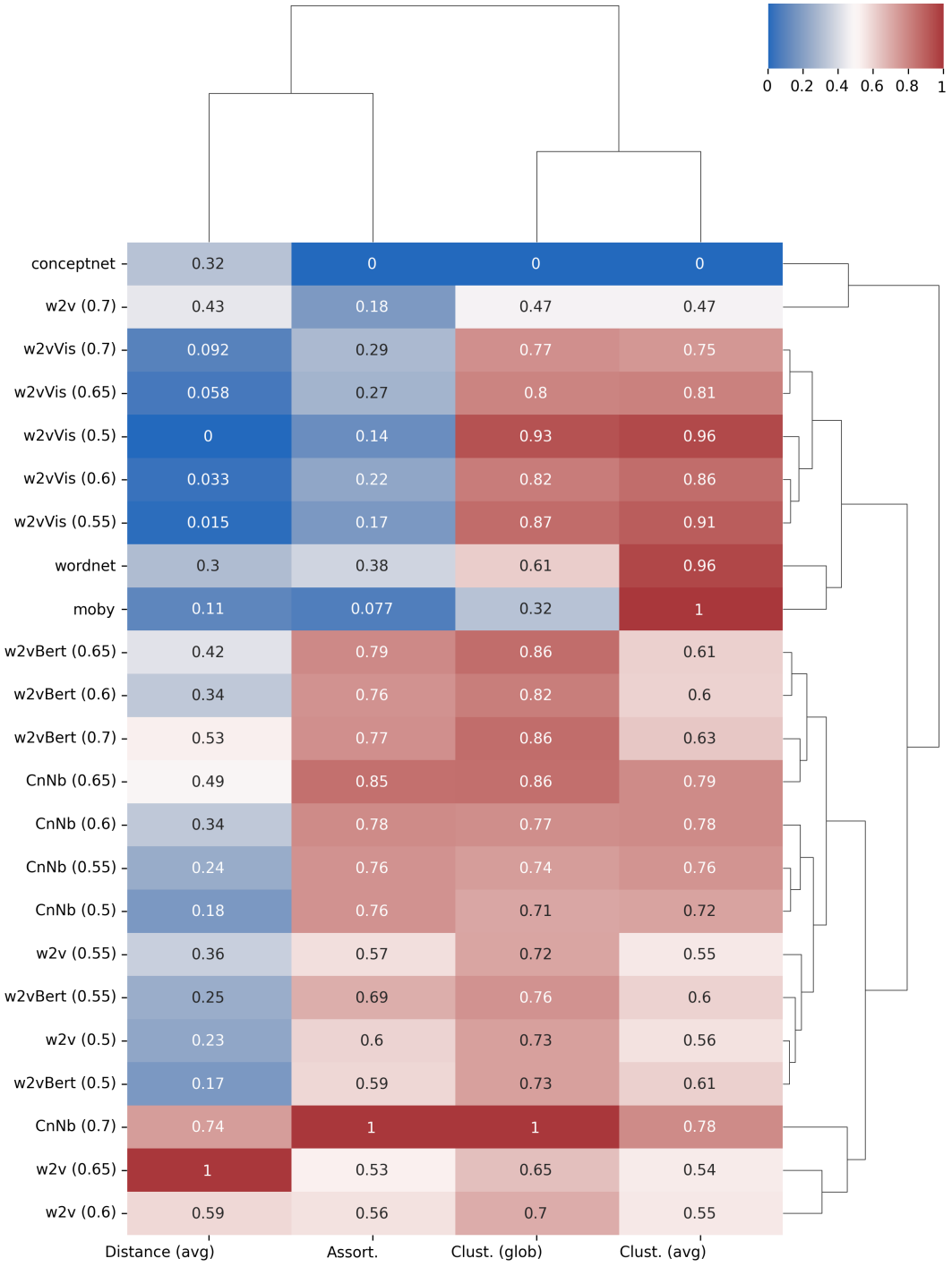
نکته جالب توجه دیگر این است که روند تغییرات مشابهی که با افزایش آستانه تشابه در بازنمایی VisWord2Vec مشاهده کردیم، نشان می‌دهد که گراف‌های این بازنمایی علی رغم تفاوت در تعداد یال‌هایشان، به یکدیگر شبیه‌اند. این موضوع می‌تواند بیانگر آن باشد که با حذف اتصالات محدود از این بازنمایی، ساختار آن تغییر زیادی نمی‌کند. در مقابل این بازنمایی، بازنمایی Word2Vec قرار دارد که با توجه به نتایج خوشه‌بندی نیز، شباهت چندانی به یکدیگر ندارند. لذا می‌توان گفت حذف تعداد محدودی از اتصالات گراف این بازنمایی، منجر به تغییرات زیادی در خواص آن شده است.

الگوی متفاوت خواص بازنمایی Word2Vec در آستانه ۰٫۷ که در بخش‌های قبلی به آن اشاره شد، در این خوشه‌بندی نیز مشهود است. گراف مذکور با کانسپت‌نت در یک خوشه قرار گرفته است. این در حالی است که سایر گراف‌های این بازنمایی با معماری‌های دیگر توسعه یافته Word2Vec در یک خوشه قرار گرفته‌اند. با این تفاسیر، می‌توان نتیجه گرفت که این بازنمایی در آستانه تشابه بالا، رفتار کاملا متفاوتی دارد.

یکی دیگر از موارد قابل بحث در این خوشه‌بندی، نزدیک‌تر بودن گراف‌های حاصل از VisWord2Vec به

[9] Agglomerative Hierarchical Clustering
[10] Min-max Scaling

گراف‌های مبتنی بر دانش انسانی نسبت به گراف‌های مبتنی بر Word2Vec است. ممکن است افزودن اطلاعات تصویری منجر به پدید آمدن یک بازنمایی نزدیک‌تر به بازنمایی‌های ساخته شده توسط انسان شده باشد.



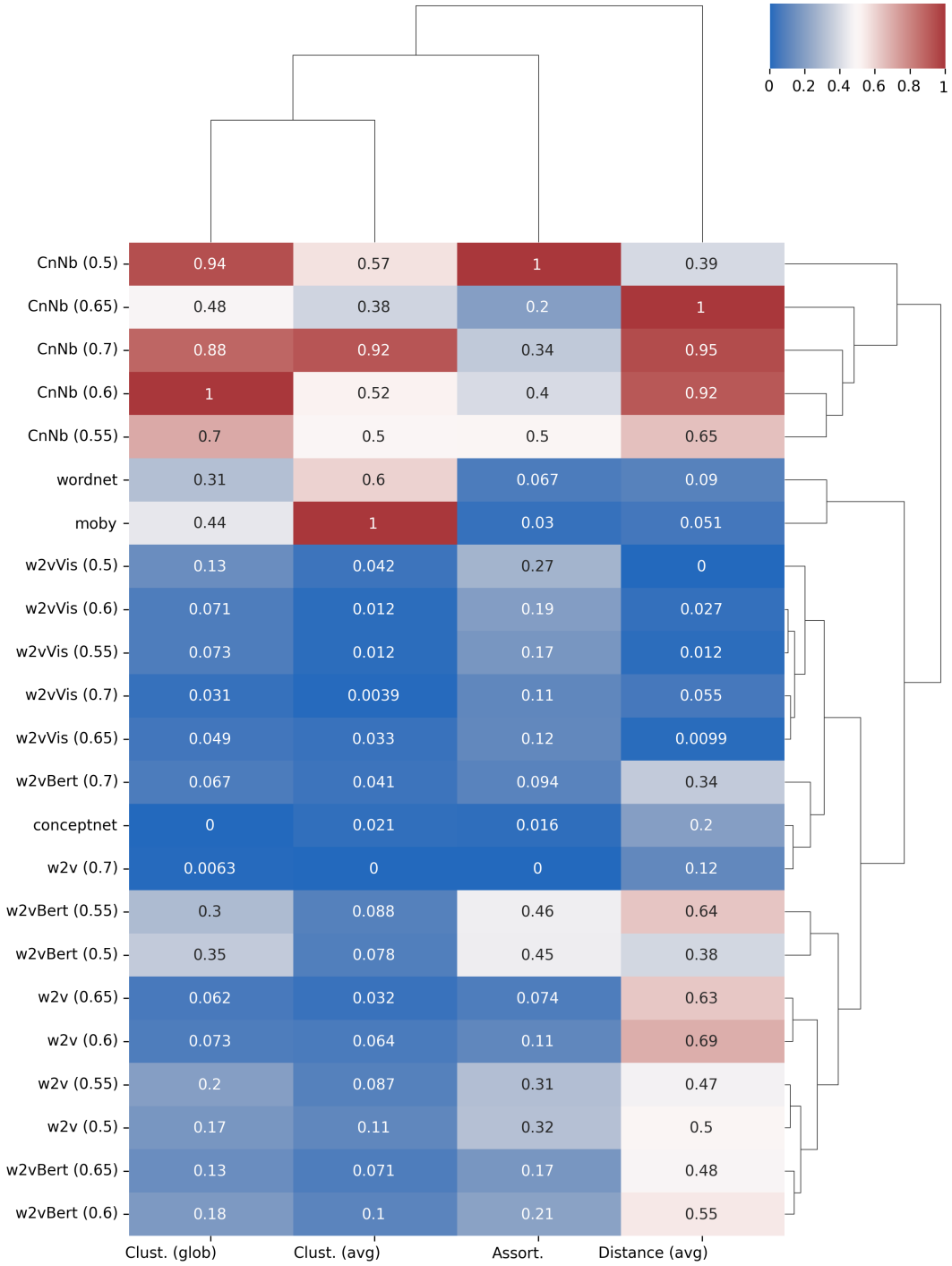
شکل ۳.۴: ماتریس رنگی نمایانگر مقادیر شاخص‌های سراسری در گراف‌های معنایی. در این تصویر، سطرها نماینده گراف‌ها و ستون‌ها (به ترتیب از راست به چپ) نماینده چهار شاخص سراسری متوسط ضریب خوشه‌بندی محلی، ضریب خوشه‌بندی سراسری، ضریب همسان‌گرایی درجه‌ای و متوسط فاصله هستند. مقادیر گزارش شده در این تصویر همگی نرمال شده‌اند. در طیف رنگی حاضر، سلول‌های آبی مقادیر کمتر و سلول‌های قرمز مقادیر بیشتری برای هر شاخص دارند. نتایج خوشه‌بندی گراف‌های معنایی در سمت راست تصویر، و نتایج خوشه‌بندی شاخص‌ها در سمت بالا مشخص است.

به طور کلی، مقایسه مستقیم مقادیر به دست آمده برای شاخص‌های مختلف میان گراف‌های معنایی ممکن است در فهم شباهت‌ها و تفاوت‌های آن‌ها چندان راه‌گشا نباشد. این محدودیت از آن‌جا ناشی می‌شود که این گراف‌ها تعداد رئوس و تعداد یال‌های متفاوتی دارند و مقایسه مستقیم نتایج شاخص‌ها ممکن است گمراه کننده باشد. به همین علت، در بخش بعدی به مقایسه نتایج معیارهای سراسری و آزمون آماری این نتایج می‌پردازیم.

۴.۳.۴ طبقه‌بندی گراف‌های معنایی با استفاده از مقادیر آزمون آماری

در بخش قبلی، تمامی مقادیر به دست آمده برای شاخص‌های سراسری در بیست و سه گراف مورد مطالعه ارائه و بررسی شد. همان‌طور که در بخش‌های گذشته ذکر شد، با توجه به غیر هم‌اندازه بودن گراف‌های مورد مطالعه، مقایسه مستقیم برخی شاخص‌های سراسری مشکل است و اغلب گویای تمامی تفاوت‌های میان گراف‌ها نیست. به بیان دیگر، شباهت مقدار شاخص‌ها برای دو گراف غیرهم‌اندازه لزوما نشان‌دهنده شباهت ساختاری آن‌ها نیست. از این رو، در این پژوهش چهارچوب نوینی برای مقایسه بهتر گراف‌های غیرهم‌اندازه ارائه شد. طبق توضیحات پیشین، یادآوری می‌شود که برای انجام این مقایسه غیر مستقیم، ابتدا هر گراف، با مدل‌های پیکربندی تصادفی همتای خودش مقایسه می‌شود. در این مقایسه، معیار ما، میزان تفاوتی است که هر گراف در هر شاخص سراسری با مدل‌های تصادفی همتای خودش دارد. پس از انجام این مقایسه، نتایج حاصل از آن برای مقایسه گراف‌های مورد مطالعه به کار می‌رود. در ادامه نتایج مربوط به این مقایسه آماری ارائه شده و طبق آن به طبقه‌بندی گراف‌های معنایی می‌پردازیم. همچنین نتایج این بخش با بخش پیشین نیز مقایسه شده است تا تفاوت این روش مقایسه غیرمستقیم با روش ارائه شده در بخش قبلی برای خواننده مستدل شود.

همان‌طور که در فصل سوم گفته شد برای مقایسه شبکه‌های معنایی غیرهم‌اندازه ابتدا ده مدل پیکربندی تصادفی هر شبکه معنایی ساخته شد. سپس چهار شاخص سراسری متوسط فاصله، ضریب همسانگرایی درجه‌ای، و ضرایب خوشه‌بندی محلی و سراسری برای هر شبکه و مدل‌های تصادفی متناظر آن اندازه‌گیری شد. با توجه به توزیع نرمال نتایج، با استفاده از آزمون آماری t به مقایسه هر گراف معنایی با مدل‌های تصادفی خودش پرداختیم. پس از محاسبه مقادیر احتمالی، دریافتیم که تمامی شبکه‌های معنایی مورد مطالعه در تحقیق جاری در تمامی شاخص‌ها، به طرز معناداری از مدل‌های تصادفی متناظرشان متفاوت هستند (مقدار احتمالی بسیار کوچک و کمتر از ۰/۰۱).



شکل ۴.۴: ماتریس رنگی نمایان‌گر مقادیر اندازه تاثیر حاصل از مقایسه شبکه‌های معنایی با مدل‌های پیکربندی تصادفی آن‌ها. در این تصویر (از راست به چپ) ستون‌ها نماینده چهار شاخص سراسری متوسط فاصله، ضریب همسان‌گرایی درجه‌ای، متوسط ضریب خوشه‌بندی محلی و ضریب خوشه‌بندی سراسری هستند. هر سطر، نماینده یکی از گراف‌های معنایی است. هر چه مقادیر یک خانه بیشتر باشد، به معنای بیشتر بودن اندازه تاثیر و بیشتر بودن تفاوت میان هر گراف معنایی با مدل‌های پیکربندی تصادفی خودش، در شاخص مربوطه است.

در گام بعدی، به منظور مقایسه شبکه‌های معنایی با یکدیگر، مقادیر اندازه تاثیر را مطابق آن‌چه در فصل سوم گفته شد محاسبه کردیم. این مقادیر نشان می‌دهند که هر گراف در هر شاخص تا چه حد از نمونه‌های تصادفی متناظر خودش متفاوت است. با این‌کار، از آنجایی که توزیع نهایی برای مقایسه تمامی شبکه‌ها به یک توزیع استاندارد نگاشت شده است به یک زمینه مشترک دست یافته‌ایم که می‌توانیم از آن برای مقایسه شبکه‌های معنایی استفاده کنیم. مقادیر مربوط به اندازه تاثیر در تصویر ۴.۴ آمده است.

همان‌طور که از تصویر برمی‌آید روند کلی که در ماتریس پیشین مشاهده شده در این ماتریس نیز قابل مشاهده است. برای مثال، بازنمایی‌های کانسپت‌نت و موبی که ضرایب همسان‌گرایی بسیار کوچکی داشتند، تفاوت کمی با گراف‌های تصادفی خودشان در این شاخص دارند. اما در ماتریس جدید که بر مبنای اندازه تاثیر ساخته شده، تفاوت‌هایی حاصل شده که به آن می‌پردازیم. در خوشه‌بندی جدید می‌توان سه دسته کلی مشاهده کرد. از طرف دیگر می‌توان دید که برخلاف آنچه در تصویر ۳.۴ مشهود بود، در اینجا خوشه‌بندی‌ها به شکل منسجم‌تری انجام شده است، چرا که گراف‌هایی که روش ساخت مشابهی دارند عمدتا در خوشه‌های مشترکی قرار گرفته‌اند. با این توضیح، می‌توان دید که در خوشه‌بندی جدید که بر اساس مقادیر اندازه تاثیر انجام شده، گراف‌های مربوط به بازنمایی Conceptnet Numberbatch همگی در یک دسته قرار گرفته‌اند. اغلب گراف‌های دو بازنمایی Word2Vec نیز در خوشه‌های مشابهی جای گرفته‌اند. تنها در آستانه ۰٫۷ گراف این بازنمایی در خوشه متفاوتی جا گرفته است. همانطور که در گزارش جدول اول و ۳.۴ گفته شد، در این آستانه تشابه، رفتار گراف بازنمایی Word2Vec به کلی تغییر می‌کند. این موضوع بار دیگر در این خوشه‌بندی مشهود است. این در حالی است که با افزایش آستانه تشابه در بازنمایی‌های دیگر چنین الگویی دیده نمی‌شود.

نتیجه دیگر که از خوشه‌بندی جدید به دست می‌آید، شباهت دو بازنمایی کانسپت‌نت و VisWord2Vec است. می‌دانیم که بازنمایی اول مبتنی بر دانش انسانی است و بازنمایی دوم مبتنی بر یادگیری ماشین و به دست آمده از یک فضای برداری است. با وجود این تفاوت‌ها، قرار گرفتن آن‌ها در یک خوشه‌بندی از این جهت قابل توجه است که هر دو این بازنمایی‌های برای بهبود کارایی در بازشناسی عبارات مبتنی بر فهم متعارف عملکرد خوبی دارند. برای مثال، با ترکیب بازنمایی کانسپت‌نت با مدل‌های زبانی، عملکرد آن‌ها در این امر به طرز چشم‌گیری افزایش پیدا می‌کند [۳۰]. از طرفی، آموزش یک مدل بر داده تصویری علاوه بر داده متنی نیز عملکرد مدل‌ها را بهبود می‌دهد [۲۰].

یکی دیگر از مشاهدات حاصل از این خوشه‌بندی تازه، تفاوت بازنمایی Conceptnet Numberbatch از سایر بازنمایی‌های معنایی است که در خوشه‌بندی قبلی که با استفاده از مقادیر حاصل از اندازه‌گیری شاخص‌ها به

دست آمده بود، مشهود نبود. همان‌گونه که می‌توان دید، این بازنمایی، در اغلب شاخص‌ها بیشترین میزان تفاوت را با مدل‌های تصادفی متناظرش دارد. شاید این امر ناشی از ترکیب چند بازنمایی مختلف (از جمله کانسپت‌نت، Word2Vec و FastText) برای ساخت این بازنمایی باشد.

نزدیک‌ترین این شبکه‌های معنایی به مدل‌های تصادفی، گراف کانسپت‌نت، و گراف‌های با آستانه تشابه ۰٫۷ در دو بازنمایی BERT2Static و Word2Vec می‌باشند.

۵.۳.۴ همپوشانی گره‌های تاثیرگذار شبکه‌های معنایی با کلمات پرتکرار زبان انگلیسی

یکی از موضوعاتی که پرداختن به آن هنگام مقایسه بازنمایی‌های معنایی ضروری است، چگونگی بازنمایی کلمات پرکاربرد در زبان است. از آنجایی که گویشوران یک زبان، چنین کلماتی را به طور مکرر استفاده میکنند، انتظار می‌رود که کلمات هم‌معنا یا مرتبط زیادی با آن کلمات در زبان وجود داشته باشد. انتظار داریم که در شبکه‌های معنایی نیز این پدیده را به نحوی مشاهده کنیم. به همین منظور، با استفاده از شاخص‌های مرکزیت، به بررسی گره‌های تاثیرگذار در شبکه‌های معنایی می‌پردازیم. همانطور که در فصل پیشین اشاره شد، در این پژوهش به منظور شناسایی مفاهیم مهم در بازنمایی‌های معنایی، از دو شاخص مرکزیت بر حسب درجه و مرکزیت بر حسب الگوریتم پیج‌رنک استفاده شده است. همچنین نسبت متوسط درجه کلمات پرتکرار به کل کلمات در این بازنمایی‌ها نیز مورد بررسی قرار گرفته و نتایج آن ارائه شده است.

جدول ۶.۴: مقایسه متوسط درجات برای کل کلمات و نسبت متوسط درجات برای کلمات پرتکرار زبان انگلیسی به کل کلمات در بازنمایی‌های معنایی

بازنمایی معنایی	تمامی کلمات	۱۰۰۰ کلمه پرتکرار	۲۰۰۰ کلمه پرتکرار	۳۰۰۰ کلمه پرتکرار
وردنت	۷٫۳۳	۲٫۷۶	۲٫۵۳	۲٫۴
موبی	۳۳٫۳۴	۶٫۰	۵٫۶۶	۵٫۲۸
کانسپت‌نت	۲٫۲۴	۲۵٫۷۵	۲۱٫۰	۱۷٫۷۹
Word2Vec	۶۹٫۸۹	۰٫۱۲	۰٫۱۵	۰٫۱۶
Word2Vec دیداری	۱۳۵۲٫۶	۰٫۷۶	۰٫۷۹	۰٫۸۲
BERT ایستا	۱۱۹٫۸۸	۰٫۰۳	۰٫۰۳	۰٫۰۳
کانسپت‌نت نامبربچ	۷۵٫۵۵	۰٫۴۱	۰٫۴۵	۰٫۴۶

به منظور مطالعه کمی این موضوع، در جدول ۶.۴ نسبت متوسط درجات ۱۰۰۰، ۲۰۰۰ و ۳۰۰۰ پرتکرارترین

کلمات زبان انگلیسی به متوسط درجات تمامی کلمات در هر یک از بازنمایی‌های معنایی نشان داده شده است. آنگونه که از جدول مشهود است، به طور کلی در بازنمایی‌های مبتنی بر دانش انسانی، کلمات پرتکرار درصد قابل توجهی از روابط معنایی را به خود اختصاص می‌دهند. این در حالی است که در تمام بازنمایی‌های مبتنی بر مدل‌سازی توزیعی معنا، کلمات پرتکرار سهم بسیار ناچیزی (کمتر از ۱) از کل روابط دارند. با این وجود در این جدول، Word2Vec دیداری نسبت به سایر بازنمایی‌های توزیعی اعداد بالاتری به خود اختصاص داده است.

برای اینکه نتایج کمی جدول ۶.۴ ملموس‌تر شود، به کلماتی که بالاترین امتیاز مرکزیت را برحسب دو معیار درجه و پیج‌رنک کسب کرده‌اند، نگاهی می‌کنیم. دو جدول ۷.۴ و ۸.۴، به ترتیب پانزده کلمه با بیشترین امتیاز مرکزیت را در بازنمایی‌های مبتنی بر دانش انسانی و بازنمایی‌های توزیعی نشان می‌دهند.

جدول ۷.۴: تاثیرگذارترین گره‌ها (مفاهیم) در شبکه‌های معنایی مبتنی بر پایگاه داده واژگان

	وردنت		موبی		کانسپت‌نت	
رتبه	درجه	پیج‌رنک	درجه	پیج‌رنک	درجه	پیج‌رنک
۱	pass	pass	cut	language	person	person
۲	break	break	set	cheese	people	people
۳	hold	get	turn	english	poly	poly
۴	check	take	run	magpie	house	water
۵	take	make	line	color	water	house
۶	run	hold	check	wine	work	eous
۷	get	check	break	pigment	head	cat
۸	go	go	color	fish	snow	sex
۹	line	run	pass	cut	back	snow
۱۰	cut	deal	light	philosopher	eous	work
۱۱	charge	see	point	parts	cat	human
۱۲	make	beat	close	silver	fish	head
۱۳	passing	set	flat	set	child	logy
۱۴	set	passing	charge	device	sex	dog
۱۵	draw	cut	cast	turn	man	child

با نظر به جدول ۷.۴، می‌توان گفت رئوس با مرکزیت بیشتر، همپوشانی خوبی با کلمات پرتکرار زبان انگلیسی دارند. از طرفی، اگر به جدول ۸.۴ که مربوط به بازنمایی‌های مبتنی بر مدل‌سازی توزیعی است نگاه کنیم در می‌یابیم که در تمامی آن‌ها، پانزده کلمه‌ای که بالاترین امتیاز مرکزیت را کسب کرده‌اند به کلمات معمول زبان انگلیسی نیستند و عموما کلمات بسیار نادر و یا اسامی خاص می‌باشند. تنها بازنمایی که در جدول ۸.۴ کلمات معمول بیشتری را در آن می‌توان مشاهده کرد، بازنمایی VisWord2Vec است.

جدول ۸.۴: تاثیرگذارترین گره‌ها (مفاهیم) در شبکه‌های ساخته شده از فضاهای معنایی مبتنی بر مدل‌سازی توزیعی

	Word2Vec		BERT2Static		VisWord2Vec		CnNumberbatch	
رتبه	درجه	پیچ‌رنک	درجه	پیچ‌رنک	درجه	پیچ‌رنک	درجه	پیچ‌رنک
۱	glomerular	animality	megacolon	superciliousness	engineering	engineering	travellerś tree	in_circles
۲	leiomyoma	pompousness	ectasia	meretricious	thirst	thirst	rose_globe_lily	flowering_plant
۳	lichen_planus	archness	proctitis	pomposity	feminist	boil	white_globe_lily	boron_counter_tube
۴	eccrine	lumpish	atrophic	mawkish	Truman	adhesive	yellow_globe_lily	philosophical_doctrine
۵	peroxidase	abjection	polymyositis	sickeningly	adhesive	reasonable	globe_lily	chemical_compound
۶	pyrimidine	rebarbative	hepatomegaly	pretentiousness	reasonable	boil	mushroom_pimple	shore_bird
۷	umbilical_vein	superciliousness	nephrolithiasis	conceptuality	feminist	Truman	flowering_plant	pelvic_inflammatory_disease
۸	thyrotropin	lecher	achlorhydria	sententious	wholesale	goblet	boron_counter_tube	travellers_tree
۹	adrenal_cortex	ruefulness	emphysematous	moronic	surgical	liking	bignoniad	chicken_cordon_bleu
۱۰	intraventricular	monism	suppurative	humorless	Ms.	enlarge	lipstick_plant	edible_fruit
۱۱	hamartoma	Mozartian	leukopenia	tiresomely	jewellery	surgical	water_carpet	anatomical_structure
۱۲	mesenteric	ideational	pachysandra	crassness	ruffle	high_ceilinged	psilophyton	viral_infection
۱۳	griseofulvin	teleology	cholelithiasis	laughably	website	glue	chemical_compound	mushroom_pimple
۱۴	mucinous	denotative	parotitis	anti-intellectual	TRUE	designing	broomweed	sinusoidal_projection
۱۵	intracerebral	positivistic	theophylline	guileless	value	website	singletary_pea	bignoniad

۴.۴ ویژگی‌های مقیاس میانی شبکه‌های معنایی

در این بخش، به بررسی بازنمایی‌های معنایی مورد مطالعه در سطح میانی می‌پردازیم. به این منظور، در هر گراف معنایی ابتدا جوامع آن شناسایی می‌شوند. سپس، با نظر به این جوامع، امتیاز پودمانگی[11] برای هر گراف معنایی گزارش شده است. نمونه‌ای از این جوامع نیز در ادامه ارائه شده است. در جدول ۹.۴ مقادیر محاسبه شده امتیاز پودمانگی برای هر یک از گراف‌های معنایی آورده شده است. با توجه به این مقادیر، می‌توان دید که امتیاز پودمانگی در گراف‌های مربوط به بازنمایی VisWord2Vec مقداری بسیار کوچک داشته و از این نظر تفاوت زیادی با سایر بازنمایی‌ها دارد. بزرگترین امتیاز پودمانگی نیز برای بازنمایی Conceptnet Numberbatch مشاهده شد. با توجه به این نتایج، می‌توان گفت که روابط معنایی که به واسطه در نظر گرفتن دادگان تصویری در VisWord2Vec به بازنمایی Word2Vec اضافه شده‌اند، عمدتا روابطی میان جوامع در این بازنمایی بوده‌اند. به بیان دیگر، لحاظ کردن اطلاعات معنایی تصویری اتصالات میان-پودمانی را افزایش و احتمال تشکیل جوامع متراکم را کاهش داده است. در بازنمایی Conceptnet Numberbatch ترکیب چند نوع بازنمایی پایه با هم، سبب افزایش احتمال تشکیل جوامع شده است.

[11] Modularity Score

جدول ۹.۴: امتیاز پودمانگی در بازنمایی‌های معنایی

بازنمایی معنایی	امتیاز پودمانگی
وردنت	۰٫۸۲
موبی	۰٫۵۷
کانسپت‌نت	۰٫۸۸
Word2Vec (۰٫۵)	۰٫۶۸
BERT2Static (۰٫۵)	۰٫۶۶
VisWord2Vec (۰٫۵)	۰٫۲
Conceptnet Numberbatch (۰٫۵)	۰٫۷۹

فصل ۵

بحث و نتیجه‌گیری

در فصل‌های پیشین موضوع مطالعه و طبقه‌بندی بازنمایی‌های معنایی با بهره‌گیری از ابزارهای علم شبکه معرفی شد. به این منظور به توضیح اهمیت انجام این پژوهش و مزایای استفاده از شاخص‌های شبکه به هدف انجام مطالعه تطبیقی میان بازنمایی‌های معنایی پرداختیم. پس از معرفی انواع بازنمایی‌های معنایی و مروری کلی بر پژوهش‌های انجام شده تا امروز، محدودیت مطالعات پیشین را شناسایی کردیم. با توجه به این مطالب، در نهایت چهارچوب پیشنهادی این پژوهش به منظور مقایسه بازنمایی‌های مختلف معنایی معرفی شد. در فصل چهارم نیز نتایج حاصل از به‌کارگیری این چهارچوب ارائه شدند.

این فصل، جمع‌بندی است از آنچه در چهار فصل گذشته آمده است. به علاوه دستاوردهای این پژوهش در این قسمت بحث خواهد شد. هم‌چنین محدودیت‌های نظری و عملی که در چهارچوب پیشنهادی با آن روبه‌رو بودیم را در این قسمت مطرح می‌کنیم. با توجه به مطالعات اندک انجام شده در طبقه‌بندی بازنمایی‌های معنایی و هم‌چنین نبود چهارچوبی مشخص جهت طبقه‌بندی گراف‌های معنایی غیرهم‌اندازه، مسیرهای متعددی برای گسترش این پژوهش می‌توان متصور شد. به همین علت، این فصل را با ارائه کاربردهای احتمالی، پیشنهادها و راه‌کارهایی برای پژوهش بیشتر در این زمینه تمام می‌کنیم.

۱.۵ جمع‌بندی

خلاصه آنچه در فصول گذشته آمده است:

- در فصل اول، موضوع مطالعه و طبقه‌بندی بازنمایی‌های معنایی معرفی شد. همچنین، اهمیت انجام این پژوهش از جنبه‌ها مختلف بررسی شد. همان‌طور که گفته شد، طبقه‌بندی بازنمایی‌های معنایی امکان مقایسه روش ساخت آن‌ها و هم‌چنین تفسیر بهتر چگونگی عمل‌کرد آن‌ها را فراهم می‌کند. به علاوه، با شناخت بهتر ویژگی‌های ساختاری بازنمایی‌های معنایی می‌توانیم از این دانش در جهت ساخت و یا طراحی بازنمایی‌های معنایی در پردازش زبان طبیعی بهره بگیریم.

- در فصل دوم با مرور پژوهش‌های انجام شده در موضوع مطالعه ساختاری و طبقه‌بندی بازنمایی‌های معنایی، محدودیت‌های موجود شناسایی ومعرفی شدند. به رغم آنکه پژوهش‌های بی‌شماری به منظور ارزیابی عملکرد بازنمایی‌های معنایی در وظایف مختلف زبانی صورت گرفته، تعداد پژوهش‌هایی که به بررسی ساختار معنایی این بازنمایی‌ها بپردازند بسیار کمتر هستند. هم‌چنین مطالعات صورت گرفته عمدتا به مقایسه بازنمایی‌های توزیعی با یک‌دیگر پرداخته و مقایسه با بازنمایی‌های مبتنی بر دانش انسانی کم‌تر مورد توجه قرار گرفته است. به علاوه، پژوهش‌های انگشت‌شماری وجود دارند که به رغم امکانات مفیدی که علم شبکه به دست می‌دهد، از این معیارها به منظور مطالعه بازنمایی‌های معنایی استفاده کنند. این در حالی است که در سایر رویکردهای مطالعه بازنمایی‌های معنایی نیز عمده روش‌ها به بررسی جنبه‌های محدودی از این بازنمایی‌ها می‌پردازند.

- در فصل سوم به توضیح رویکرد اتخاذ شده به منظور مطالعه و مقایسه بازنمایی‌های معنایی پرداختیم. ابتدا چگونگی نگاشت بازنمایی‌های معنایی مورد مطالعه به گراف متناظرشان شرح داده شد. از آنجایی که بازنمایی‌های اولیه در دو دسته کلی انسانی و توزیعی جا می‌گیرند، مراحل نگاشت هر گروه به طور مجزا توضیح داده شد. پس از دست‌یابی به بازنمایی شبکه‌ای برای تمامی بازنمایی‌های اولیه، به ارائه چهارچوب پیشنهادی مقایسه شبکه‌های معنایی حاصل پرداختیم. این چهارچوب، مبتنی بر شاخص‌های علم شبکه است و دو بخش شاخص‌های مقیاس سراسری و شاخص‌های مقیاس میانی را در بر می‌گیرد. در بخش شاخص‌های مقیاس سراسری، به معرفی رویکرد جدید آماری جهت طبقه‌بندی بازنمایی‌ها معنایی غیرهم‌اندازه پرداختیم.

- با توجه به تعدد شاخص‌های مورد استفاده یافته‌های این پژوهش را به صورت موردی بیان می‌کنیم. در این قسمت، ارتباط یافته‌های فصل چهارم با مطالعات پیشین بررسی می‌شود.

- با مطالعه ساختاری بازنمایی‌های معنایی انسانی دریافتیم که بر خلاف آن‌چه قبلا درباره این بازنمایی‌ها نتیجه‌گیری شده بود [۱۰]، مبتنی بر دانش انسانی بودن لزوما موجب ساختار یکسان در بازنمایی معنایی نمی‌شود. متوجه شدیم که برخلاف دو بازنمایی انسانی وردنت و موبی، بازنمایی کانسپت‌نت گرافی بسیار تنک است، ضرایب خوشگی بسیار کوچک و متوسط فاصله نسبتا بزرگی دارد. لذا برخلاف آن‌چه ورمیف و همکاران[1] در مورد دو بازنمایی انسانی موبی و وردنت مطرح می‌کنند، بازنمایی کانسپت‌نت را نمی‌توان جهان‌کوچک در نظر گرفت.

- از میان بازنمایی‌های مطالعه شده، گراف‌های بازنمایی VisWord2Vec بیش‌ترین میزان تراکم معنایی و کمترین مقدار متوسط فاصله را دارند. این موضوع نشان می‌دهد که این بازنمایی نسبت قابل توجهی از روابط معنایی بالقوه را در نظر گرفته است. از طرفی، می‌دانیم که این بازنمایی، تنها بازنمایی مورد مطالعه است که از اطلاعات تصویری جهت استخراج روابط معنایی بهره می‌گیرد. لذا می‌توان این‌طور نتیجه گرفت که با استفاده از اطلاعات معنایی موجود در دادگان تصویری می‌توان به روابط معنایی قابل توجهی دست پیدا کرد که لزوما در دادگان متنی قابل دسترس نیست. این مشاهده با ادعای بندر و کولر نیز همسو می‌باشد. آن‌ها بیان می‌کنند که مدل‌های زبانی که تنها بر داده متنی آموزش داده شده‌اند لزوما به تمام روابط معنایی ممکن دسترسی ندارند. به عقیده آن‌ها استفاده از دادگان حسی برای آموزش مدل‌های زبانی امکان دسترسی آن‌ها به روابط معنایی بیشتری را فراهم می‌کند.

- در بازنمایی VisWord2Vec، با افزایش آستانه تشابه تغییر چندانی در ویژگی‌های سراسری گراف حاصل از این بازنمایی رخ نمی‌دهد. این در حالی است که با افزایش آستانه تشابه در بازنمایی Word2Vec و BERT2Static، ساختار گراف حاصل تغییرات زیادی نشان می‌دهد.

- با نظر به میزان تراکم یال در بازنمایی‌های معنایی، می‌توان مشاهده کرد که کمترین میزان این شاخص مربوط به کانسپت‌نت، وردنت و موبی است. نقطه اشتراک تمامی این بازنمایی‌ها این است که منبع دادگان آن‌ها به طور کامل دانش انسانی بوده است. با توجه به محدودیت‌های روش ساخت با استفاده از دانش انسانی، می‌توان گفت که روابط معنایی زیادی نادیده گرفته شده‌اند.

- پس از تصویرسازی توزیع درجات بازنمایی‌های معنایی، مشاهده کردیم که تمامی آن‌ها، اعم از

[1] Veremyev et al. 2019

انسانی و توزیعی، انحراف مثبت دارند. این مشاهده به این معناست که در تمامی بازنمایی‌های معنایی، تعداد بسیار محدودی از کلمات ارتباطات معنایی زیادی دارند.

- با استفاده از معیارهای مرکزیت و مقایسه متوسط درجات کلمات پرتکرار زبان انگلیسی در هر یک از بازنمایی‌ها، دریافتیم که در بازنمایی‌های مبتنی بر دانش انسانی، کلمات پرتکرار جایگاه موثرتری به خود اختصاص می‌دهند. این در حالی است که در تمامی بازنمایی‌های توزیعی، نه تنها کلمات پرتکرار رئوس تاثیرگذار نیستند، بلکه کلمات بسیار نادری میان رئوس تاثیرگذار با متوسط درجه بالا قرار می‌گیرند. با توجه به فرضیه توزیعی، انتظار داشتیم که در بازنمایی‌های توزیعی نیز کلمات پرتکرار روابط معنایی بیشتری داشته باشند اما عکس این موضوع مشاهده شد. دلایل این امر می‌تواند مورد پژوهش بیشتر واقع شود.
- مشاهده کردیم که استفاده از مقادیر اندازه تاثیر به منظور طبقه‌بندی گراف‌های معنایی غیرهم‌اندازه نتایج منسجم‌تری به دست می‌دهد. از جمله اینکه خوشه‌بندی دقیق‌تری صورت می‌گیرد و گراف‌های متعلق به یک بازنمایی در خوشه‌های یکسان قرار می‌گیرند.
- با مطالعه گراف‌های معنایی در سطح میانی، اطلاعات بیشتری درباره بازنمایی‌های معنایی حاصل شد. با محاسبه امتیاز پودمانگی در گراف‌های معنایی دریافتیم که در بازنمایی VisWord2Vec احتمال یافتن جوامع متراکم بسیار پایین‌تر از سایر بازنمایی‌هاست. با توجه به روش ساخت این بازنمایی، می‌توان گفت روابط معنایی به‌دست‌آمده از اطلاعات تصویری، عمدتا میان جوامع معنایی موجود در Word2Vec برقرار شده و به همین علت احتمال تشکیل جامعه در بازنمایی نهایی بسیار کم شده است.

۲.۵ نوآوری

در این بخش به بررسی نوآوری‌ها و مواردی که در این پژوهش برای اولین بار انجام شده‌اند می‌پردازیم.

در این پژوهش هفت نوع بازنمایی معنایی توزیعی و انسانی مورد مطالعه و طبقه‌بندی قرار گرفته‌اند که پیش از این، مطالعه‌ای در این مقیاس از جامعیت انجام نشده بود. این پژوهش، با توجه به اهمیت مطالعه و مقایسه جامع بازنمایی‌های معنایی که امروزه در امور مختلف پردازش زبان طبیعی کاربرد دارند، انجام گرفت. اولین نوع

بازنمایی که نخستین بار در این پژوهش مورد مطالعه ساختاری قرار گرفته است، بازنمایی حاصل از آموزش یک مدل شبکه عصبی بر داده چندوجهی متن-تصویر است. از آنجایی که در سال‌های اخیر استفاده از اطلاعات غیر متنی (نظیر تصویر و صوت) به منظور غنی کردن بازنمایی‌های معنایی متنی مورد توجه بسیاری قرار گرفته بررسی این بازنمایی و تفاوت‌های آن با سایر بازنمایی‌های معمول ضروری می‌نمود. نوع دوم بازنمایی که در این پژوهش برای نخستین بار به منظور انجام مقایسه ساختاری با سایر بازنمایی‌های معمول در نظر گرفته شد، ترکیب اطلاعات خارجی از بازنمایی معنایی با مدل‌های توزیعی معنا است.

حوزه دیگر نوآوری این پژوهش، مربوط به چهارچوب مقایسه شبکه‌های معنایی است. از آنجایی که شبکه‌های حاصل از بازنمایی‌های معنایی در تعداد رئوس (کلمات) و یال‌ها (روابط معنایی موجود در بازنمایی) با یکدیگر یکسان نبودند، نیاز به روشی وجود داشت که بتواند مقایسه شبکه‌های غیرهم‌اندازه معنایی را به دست دهد. به همین منظور، در این پژوهش یک چهارچوب آماری به منظور مقایسه این نوع شبکه‌ها ارائه شده است. در این چهارچوب، ابتدا هر شبکه با مدل‌های پیکربندی تصادفی همتای خودش، با یک آزمون آماری مقایسه می‌شود و نتایج این آزمون آماری، مبنای مقایسه شبکه‌های غیرهم‌اندازه قرار می‌گیرد. طبق اطلاعات ما، این روش مقایسه شبکه‌های غیرهم‌اندازه، نخستین بار در این پژوهش مورد استفاده قرار گرفته است. از آنجایی که مقایسه شبکه‌هایی که تعداد رئوس و یال‌های متفاوتی دارند ممکن است در حوزه‌های مختلفی که از مدل‌سازی شبکه استفاده می‌کنند انجام شود (از جمله شبکه‌های مغزی، شبکه‌های زیستی، شبکه‌های اجتماعی و ...)، این چهارچوب می‌تواند در پژوهش‌های حوزه‌های دیگری به غیر از بازنمایی‌های معنایی به کار گرفته شود.

۳.۵ محدودیت‌ها

اولین محدودیت چهارچوب مورد استفاده به منظور مقایسه بازنمایی‌های معنایی، عدم امکان استفاده از آن در بازنمایی‌های معنایی غیریک‌ریخت (پویا)[۲] است. با توجه به اینکه در این چهارچوب نیاز است که هر بازنمایی تنها به یک گراف نگاشت شود، عملا امکان مقایسه بازنمایی‌های غیریک‌ریخت از دست می‌رود. این در حالی است که مدل‌های زبانی که امروزه عموما بهترین عملکرد را در میان رقیبان خود دارند، مدل‌های زبانی مبتنی بر معماری مبدل[۳] هستند که با توجه به بافت کلمات در یک جمله ممکن است بازنمایی‌های متفاوتی برای یک کلمه

[2]Non-isomorphic (Dynamic)
[3]Transformer Architecture

واحد در بافت‌های متفاوت به دست دهند. این موضوع باعث عدم امکان نگاشت یک به یک بازنمایی معنایی به گراف متناظرش می‌شود، چرا که بازنمایی اولیه یک بازنمایی ایستا نیست. با وجود این محدودیت، راه حلی که در این پژوهش اتخاذ شد، مطالعه یک بازنمایی مبتنی بر Word2Vec است که با بهره‌گیری از مدل زبانی برت بهبود یافته است. هر چند به طور کلی، از آنجایی که مقایسه این نوع‌بازنمایی‌ها می‌تواند ابعاد جدیدی از ساختار آن‌ها که منجر به پیشرفت چشم‌گیر مدل‌های زبانی شده است را روشن کند، طراحی یک چهارچوب که این نوع مقایسه را میسر سازد امری مهم و کارآمد خواهد بود.

دومین محدودیتی که بر این پژوهش وارد است، عدم آموزش مدل‌ها بر مجموعه داده‌های یکسان است. با توجه به اینکه مدل‌های استفاده شده در این پژوهش، مدل‌های زبانی ازپیش‌آموزش‌دیده هستند، و آموزش آن‌ها از صفر صورت نگرفته است، در برخی موارد دادگان مورد استفاده برای آموزش این مدل‌ها متفاوت بوده است و انتخاب آن از عهده این پژوهش خارج بوده است. دلیل عدم انجام آموزش از صفر این مدل‌ها، هزینه زمانی و منابع مورد نیاز برای آموزش هرکدام از این مدل‌هاست. در صورت امکان آموزش هر مدل بر روی داده یکسان و کنترل شرایط آموزش، می‌توان زمینه بهتری به منظور انجام مقایسه میان آن‌ها فراهم کرد.

۴.۵ پیشنهادها

در بخش پایانی این فصل به ارائه پیشنهاداتی برای گسترش پژوهش حاضر و کاربردهای احتمالی آن در آینده می‌پردازیم.

۱. اولین مسیر گسترش این پژوهش می‌تواند در جهت مطالعه بازنمایی‌های معنایی چند زبانه باشد. در پژوهش حاضر، همان‌طور که در فصل‌های دو و سه شرح داده شد، تمامی بازنمایی‌های مورد مطالعه بازنمایی‌های تک‌زبانه و انگلیسی زبان هستند و حتی به منظور مقایسه بهتر، بازنمایی کانسپت‌نت نیز که تنها بازنمایی چندزبانه این مطالعه است، به قسمت انگلیسی زبان آن تقلیل یافته است. علاوه بر کانسپت‌نت، بازنمایی‌های معنایی چند زبانه متفاوتی ارائه شده‌اند که می‌توان از آنها بهره گرفت. یکی از آن‌ها گراف معنایی بابل‌نت، یکی از بزرگ‌ترین گراف‌های معنایی چندزبانه است که می‌تواند مورد استفاده قرار گیرد. به علاوه مدل‌های زبانی چندزبانه که اخیرا معرفی شده‌اند، از جمله برت چندزبانه[۴]،

[4]Multilingual BERT

که می‌توانند در چنین مطالعه‌ای استفاده شوند.

۲. مسیر بعدی که می‌توان به منظور استفاده از پژوهش حاضر متصور شد، بهره‌گیری از چهارچوب نوین ارائه شده جهت مقایسه شبکه‌های غیرهم‌اندازه است. می‌دانیم که امروزه علم شبکه در حوزه‌های متنوعی مورد استفاده قرار می‌گیرد. از جمله این حوزه‌ها، مدل‌سازی شبکه مغزی، مدلسازی روابط موجودیت‌های زیستی، مدلسازی شبکه پروتئینی و مدل‌سازی شبکه ارتباطی اجتماعی را می‌توان نام برد. در هر یک از این حوزه‌ها و به طور کلی در هر کاربردی که از مدلسازی موجودیت‌ها و روابط میان آنها به شکل گراف بهره می‌گیرد و نیاز به مقایسه چنین گراف‌هایی وجود داشته باشد، در صورت غیرهم‌اندازه بودن گراف‌های مورد مطالعه، چهارچوب ارائه شده می‌تواند مورد استفاده قرار بگیرد. به همین علت، می‌توان گفت که کارکرد این چهارچوب تنها به شبکه‌های معنایی محدود نمی‌شود.

۳. همانطور که در بخش محدودیت‌ها گفته شد، چهارچوب حاضر و علی‌الخصوص روش مورد استفاده جهت نگاشت بازنمایی‌های برداری به گراف متناظرشان امکان بررسی بازنمایی‌های پیچیده‌تر از جمله بازنمایی‌های پویا را به دست نمی‌دهد. یکی از مهم‌ترین مسیرهای گسترش این پژوهش طراحی یک سیاست تخصیص یال است که امکان نگاشت بازنمایی‌های پویا به یک شبکه را فراهم کند و به این منظور نیاز داریم که سیاست تخصیص یال بازتعریف شود. از روش‌های احتمالی می‌تواند بازتعریف راس و یال با توجه به ساختار یک بازنمایی پویا و یا در نظر گرفتن انواع مختلف یال باشد. از طرفی، ممکن است بهره‌گیری از انواع پیچیده‌تر مدلسازی شبکه‌ای که امکان مدل‌سازی روابط پیچیده‌تری را نسبت به گراف‌های ساده به دست می‌دهند راه‌گشا باشد. برای مثال ابرگراف‌ها[5] این امکان را فراهم میکنند که یک یال واحد بتواند بیشتر از دو راس به هم متصل کند. با توجه به این ویژگی، می‌توان ارتباط بین بردارهای متفاوتی که برت برای کلمات در بافت‌های مختلف به دست می‌دهد را با استفاده از ابریال[6] مدل کرد و از گراف نهایی برای مطالعه این بازنمایی استفاده کرد. روش دیگر حل این چالش، استفاده مستقیم از بردارهای برت در دوازده لایه مختلف آن، با ورودی دادن یک مجموعه داده متنی است. آن‌گاه، بررسی شبکه حاصل در هر لایه می‌تواند مورد بررسی قرار بگیرد. برای مثال، میزان خوشه‌بندی در لایه‌های مختلف ممکن است بتواند میزان حساسیت آن لایه به بافت کلمات را نشان دهد.

[5] Hypergraphs
[6] Hyperedge

۴. همانطور که در مورد اول گفته شد، تمامی بازنمایی‌های مطالعه شده در این پژوهش صرفا دربرگیرنده زبان انگلیسی هستند. از طرفی، آزمودن چهارچوب پیشنهادی بر بازنمایی‌های زبان‌های دیگر، به ویژه زبان‌های با منابع محدود[۷]، می‌تواند نتایج جالبی به دست دهد. از جمله میزان بسط پذیری نتایج این پژوهش به زبان‌های غیر انگلیسی می‌تواند مورد بحث قرار گیرد.

۵. یکی دیگر از بخش‌های این پژوهش که تغییر آن ممکن است منجر به نتایج جدید و جالبی باشد، تاثیر انتخاب نوع سیاست تخصیص یال در نگاشت بازنمایی‌های برداری به یک گراف معنایی است. در این تحقیق، برای انجام این نگاشت، هر بازنمایی برداری را بر حسب پنج آستانه تشابه به پنج گراف مستقل نگاشت کردیم و هر کدام را مورد مطالعه قرار دادیم. از تغییراتی که می‌توان در سیاست تخصیص یال اتخاذ شده لحاظ کرد، محدود کردن تعداد و یا حذف آستانه تشابه و در عوض در نظر گرفتن مقادیر به دست آمده تشابه به عنوان وزن یال‌هاست. تغییرات ساختاری احتمالی گراف حاصل با این سیاست جدید می‌تواند قابل توجه باشد.

۶. دیگر مسیر گسترش این پژوهش می‌تواند در بخش شاخص‌های میانی صورت گیرد. در تحقیق جاری، ما به یک مطالعه پایه در این مقیاس از شبکه‌های معنایی اکتفا کردیم. این در حالی است که تغییرات متنوعی در این بخش امکان‌پذیر است که می‌تواند منجر به درک ابعاد جدیدی از تفاوت‌ها و ویژگی‌های این شبکه‌ها شود. از جمله می‌توان روش شناسایی جوامع را تغییر داد و از الگوریتم‌های مبتنی بر قدم‌زن تصادفی[۸] استفاده کرد. این دسته از الگوریتم‌های شناسایی جوامع عملکرد بسیار خوبی خصوصا در شبکه‌های بزرگ دارند.

[7] Low Resource Languages
[8] Random Walker

کتاب‌نامه

[1] Miller, George A, Beckwith, Richard, Fellbaum, Christiane, Gross, Derek, and Miller, Katherine J. Introduction to wordnet: An on-line lexical database. *International journal of lexicography*, 3(4):235–244, 1990.

[2] Speer, Robyn, Havasi, Catherine, et al. Representing general relational knowledge in conceptnet 5. in *LREC*, vol. 2012, pp. 3679–86, 2012.

[3] Noever, David and McKee, Forrest. Chatbots as problem solvers: Playing twenty questions with role reversals. *arXiv preprint arXiv:2301.01743*, 2023.

[4] Abdelghani, Rania, Wang, Yen-Hsiang, Yuan, Xingdi, Wang, Tong, Sauzéon, Hélène, and Oudeyer, Pierre-Yves. Gpt-3-driven pedagogical agents for training children's curious question-asking skills. *arXiv preprint arXiv:2211.14228*, 2022.

[5] Qian, Tracy, Xie, Andy, and Bruckmann, Camille. Sensitivity analysis on transferred neural architectures of bert and gpt-2 for financial sentiment analysis. *arXiv preprint arXiv:2207.03037*, 2022.

[6] Sezgin, Emre, Sirrianni, Joseph, Linwood, Simon L, et al. Operationalizing and implementing pretrained, large artificial intelligence linguistic models in the us health care system: Outlook of generative pretrained transformer 3 (gpt-3) as a service model. *JMIR Medical Informatics*, 10(2):e32875, 2022.

[7] Ramesh, Aditya, Pavlov, Mikhail, Goh, Gabriel, Gray, Scott, Voss, Chelsea, Radford, Alec, Chen, Mark, and Sutskever, Ilya. Zero-shot text-to-image generation, 2021.

[8] Sharples, Mike. Automated essay writing: an aied opinion. *International Journal of Artificial Intelligence in Education*, 32(4):1119–1126, 2022.

[9] Castelvecchi, Davide. Can we open the black box of ai? *Nature News*, 538(7623):20, 2016.

[10] Veremyev, Alexander, Semenov, Alexander, Pasiliao, Eduardo L, and Boginski, Vladimir. Graph-based exploration and clustering analysis of semantic spaces. *Applied Network Science*, 4(1):1–26, 2019.

[11] Emerson, Guy. What are the goals of distributional semantics? *arXiv preprint arXiv:2005.02982*, 2020.

[12] Mikolov, Tomas, Chen, Kai, Corrado, Greg, and Dean, Jeffrey. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[13] Devlin, Jacob, Chang, Ming-Wei, Lee, Kenton, and Toutanova, Kristina. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[14] Peters, Matthew E., Neumann, Mark, Iyyer, Mohit, Gardner, Matt, Clark, Christopher, Lee, Kenton, and Zettlemoyer, Luke. Deep contextualized word representations. *CoRR*, abs/1802.05365, 2018.

[15] Sharma, Aditya, Talukdar, Partha, et al. Towards understanding the geometry of knowledge graph embeddings. in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 122–131, 2018.

[16] Colombo, Pierre, Noiry, Nathan, Irurozki, Ekhine, and Clémençon, Stéphan. What are the best systems? new perspectives on nlp benchmarking. *arXiv preprint arXiv:2202.03799*, 2022.

[17] Mimno, David and Thompson, Laure. The strange geometry of skip-gram with negative sampling. in *Empirical Methods in Natural Language Processing*, 2017.

[18] Chang, Tyler A, Tu, Zhuowen, and Bergen, Benjamin K. The geometry of multilingual language model representations. *arXiv preprint arXiv:2205.10964*, 2022.

[19] Rajaee, Sara and Pilehvar, Mohammad Taher. An isotropy analysis in the multilingual bert embedding space. *arXiv preprint arXiv:2110.04504*, 2021.

[20] Kottur, Satwik, Vedantam, Ramakrishna, Moura, José MF, and Parikh, Devi. Visual word2vec (vis-w2v): Learning visually grounded word embeddings using abstract scenes. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4985–4994, 2016.

[21] Speer, Robyn and Chin, Joshua. An ensemble method to produce high-quality word embeddings (2016). *arXiv preprint arXiv:1604.01692*, 2016.

[22] Gupta, Prakhar and Jaggi, Martin. Obtaining better static word embeddings using contextual embedding models. *arXiv preprint arXiv:2106.04302*, 2021.

[23] Bender, Emily M and Koller, Alexander. Climbing towards nlu: On meaning, form, and understanding in the age of data. in *Proceedings of the 58th annual meeting of the association for computational linguistics*, pp. 5185–5198, 2020.

[24] Miller, George A. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[25] Ward, G. Moby thesaurus ii. *Project Gutenberg Literary Archive Foundation*, 2002.

[26] Jimenez, Sergio, Gonzalez, Fabio A, Gelbukh, Alexander, and Duenas, George. Word2set: Wordnet-based word representation rivaling neural word embedding for lexical similarity and sentiment analysis. *IEEE Computational Intelligence Magazine*, 14(2):41–53, 2019.

[27] AlMousa, Mohannad, Benlamri, Rachid, and Khoury, Richard. A novel word sense disambiguation approach using wordnet knowledge graph. *Computer Speech & Language*, 74:101337, 2022.

[28] Jiang, Yuncheng. Semantically-enhanced information retrieval using multiple knowledge sources. *Cluster Computing*, 23(4):2925–2944, 2020.

[29] Speer, Robyn, Chin, Joshua, and Havasi, Catherine. Conceptnet 5.5: An open multilingual graph of general knowledge. in *Thirty-first AAAI conference on artificial intelligence*, 2017.

[30] Wang, Cunxiang, Liang, Shuailong, Jin, Yili, Wang, Yilong, Zhu, Xiaodan, and Zhang, Yue. Semeval-2020 task 4: Commonsense validation and explanation. *arXiv preprint arXiv:2007.00236*, 2020.

[31] Navigli, Roberto, Blloshmi, Rexhina, and Lorenzo, Abelardo Carlos Martinez. Babelnet meaning representation: A fully semantic formalism to overcome language barriers. in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, pp. 12274–12279, 2022.

[32] Harris, Zellig Sabbetai. Distributional structure. word, 10: 146–162. reprinted in: Harris (1970), papers in structural and transformational linguistics, chapter 36. *Harris (1981), Papers on Syntax*, pp. 3–22, 1954.

[33] Foltz, Peter W. Latent semantic analysis for text-based research. *Behavior Research Methods, Instruments, & Computers*, 28(2):197–202, 1996.

[34] Pennington, Jeffrey, Socher, Richard, and Manning, Christopher D. Glove: Global vectors for word representation. in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543, 2014.

[35] Bojanowski, Piotr, Grave, Edouard, Joulin, Armand, and Mikolov, Tomas. Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5:135–146, 2017.

[36] Brown, Tom, Mann, Benjamin, Ryder, Nick, Subbiah, Melanie, Kaplan, Jared D, Dhariwal, Prafulla, Neelakantan, Arvind, Shyam, Pranav, Sastry, Girish, Askell, Amanda, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[37] Harnad, Stevan. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3):335–346, 1990.

[38] Kiela, Douwe, Bulat, Luana, and Clark, Stephen. Grounding semantics in olfactory perception. in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pp. 231–236, 2015.

[39] Kiela, Douwe and Clark, Stephen. Learning neural audio embeddings for grounding semantics in auditory perception. *Journal of Artificial Intelligence Research*, 60:1003–1030, 2017.

[40] Zhang, Yundong, Niebles, Juan Carlos, and Soto, Alvaro. Interpretable visual question answering by visual grounding from attention supervision mining. in *2019 ieee winter conference on applications of computer vision (wacv)*, pp. 349–357. IEEE, 2019.

[41] Chen, Nenglun, Pan, Xingjia, Chen, Runnan, Yang, Lei, Lin, Zhiwen, Ren, Yuqiang, Yuan, Haolei, Guo, Xiaowei, Huang, Feiyue, and Wang, Wenping. Distributed attention for grounded image captioning. in *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 1966–1975, 2021.

[42] He, Yun, Zhu, Ziwei, Zhang, Yin, Chen, Qin, and Caverlee, James. Infusing disease knowledge into bert for health question answering, medical inference and disease name recognition. *arXiv preprint arXiv:2010.03746*, 2020.

[43] Bommasani, Rishi, Davis, Kelly, and Cardie, Claire. Bert wears gloves: Distilling static embeddings from pretrained contextual representations. 2019.

[44] Ethayarajh, Kawin. How contextual are contextualized word representations? comparing the geometry of bert, elmo, and gpt-2 embeddings. *arXiv preprint arXiv:1909.00512*, 2019.

[45] Abdullah, Badr M and Klakow, Dietrich. Analyzing the representational geometry of acoustic word embeddings. *arXiv preprint arXiv:2301.03012*, 2023.

[46] Shen, Ke and Kejriwal, Mayank. A data-driven study of commonsense knowledge using the conceptnet knowledge base. *arXiv preprint arXiv:2011.14084*, 2020.

[47] Dusi, Michele, Arici, Nicola, Gerevini, Alfonso E, Putelli, Luca, and Serina, Ivan. Graphical identification of gender bias in bert with a weakly supervised approach. in *Proceedings of the Sixth Workshop on Natural Language for Artificial Intelligence (NL4AI 2022) co-located with 21th International Conference of the Italian Association for Artificial Intelligence (AI* IA 2022)*, 2022.

[48] Corrêa Jr, Edilson A, Marinho, Vanessa Q, and Amancio, Diego R. Semantic flow in language networks discriminates texts by genre and publication date. *Physica A: Statistical Mechanics and its Applications*, 557:124895, 2020.

[49] Silva, Filipi N, de Arruda, Henrique F, Costa, Luciano da F, Amancio, Diego R, et al. Accessibility and trajectory-based text characterization. *arXiv preprint arXiv:2201.06665*, 2022.

[50] Correa Jr, Edilson A, Lopes, Alneu A, and Amancio, Diego R. Word sense disambiguation: A complex network approach. *Information Sciences*, 442:103–113, 2018.

[51] Cancho, Ramon Ferrer I and Solé, Richard V. The small world of human language. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1482):2261–2265, 2001.

[52] Fukś, Henryk and Krzemiński, Mark. Topological structure of dictionary graphs. *Journal of Physics A: Mathematical and Theoretical*, 42(37):375101, 2009.

[53] Solé, Ricard V, Corominas-Murtra, Bernat, Valverde, Sergi, and Steels, Luc. Language networks: Their structure, function, and evolution. *Complexity*, 15(6):20–26, 2010.

[54] Biemann, Chris, Roos, Stefanie, and Weihe, Karsten. Quantifying semantics using complex network analysis. in *Proceedings of COLING 2012*, pp. 263–278, 2012.

[55] Yenicelik, David, Schmidt, Florian, and Kilcher, Yannic. How does bert capture semantics? a closer look at polysemous words. in *Proceedings of the Third BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, pp. 156–162, 2020.

[56] An, Yuan, Kalinowski, Alexander, and Greenberg, Jane. Clustering and network analysis for the embedding spaces of sentences and sub-sentences. in *2021 Second International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*, pp. 138–145. IEEE, 2021.

[57] Lakhzoum, Dounia, Izaute, Marie, and Ferrand, Ludovic. Semantic network analysis of abstract and concrete word associations. *arXiv preprint arXiv:2110.09096*, 2021.

[58] Tantardini, Mattia, Ieva, Francesca, Tajoli, Lucia, and Piccardi, Carlo. Comparing methods for comparing networks. *Scientific reports*, 9(1):1–19, 2019.

[59] Costa, L da F, Rodrigues, Francisco A, Travieso, Gonzalo, and Villas Boas, Paulino Ribeiro. Characterization of complex networks: A survey of measurements. *Advances in physics*, 56(1):167–242, 2007.

[60] Erdős, Paul and Rényi, Alfréd. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1):261–267, 1961.

[61] Watts, Duncan J and Strogatz, Steven H. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

[62] Barabási, Albert-László and Albert, Réka. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[63] Bird, Steven. Nltk: the natural language toolkit. in *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, pp. 69–72, 2006.

[64] Pagliardini, Matteo, Gupta, Prakhar, and Jaggi, Martin. Unsupervised learning of sentence embeddings using compositional n-gram features. *arXiv preprint arXiv:1703.02507*, 2017.

[65] Ganitkevitch, Juri, Van Durme, Benjamin, and Callison-Burch, Chris. Ppdb: The paraphrase database. in *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 758–764, 2013.

[66] Raju, T Nora, Rahana, PA, Moncy, Raichel, Ajay, Sreedarsana, and Nambiar, Sindhya K. Sentence similarity-a state of art approaches. in *2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS)*, pp. 1–6. IEEE, 2022.

[67] Bedru, Hayat Dino, Yu, Shuo, Xiao, Xinru, Zhang, Da, Wan, Liangtian, Guo, He, and Xia, Feng. Big networks: A survey. *Computer Science Review*, 37:100247, 2020.

[68] Soffer, Sara Nadiv and Vazquez, Alexei. Network clustering coefficient without degree-correlation biases. *Physical Review E*, 71(5):057101, 2005.

[69] Barabási, Albert-László. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120375, 2013.

[70] Lepley, William M and Kobrick, John L. Word usage and synonym representation in the english language. *The Journal of Abnormal and Social Psychology*, 47(2S):572, 1952.

[71] Blondel, Vincent D, Guillaume, Jean-Loup, Lambiotte, Renaud, and Lefebvre, Etienne. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, oct 2008.

[72] Clauset, Aaron, Newman, Mark EJ, and Moore, Cristopher. Finding community structure in very large networks. *Physical review E*, 70(6):066111, 2004.

[73] Newman, Mark EJ. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

Abstract

Semantic representations are essential to natural language processing since they enable language models to process the meaning of words and phrases. While both language models and human-generated semantic graphs have been shown to be effective in various natural language processing tasks, the underlying representations that drive their performance are not well understood. This study aims to profile and classify semantic representations based on the structural characteristics of their graph structures. To this end, we studied the structural differences between various semantic representations such as contextualized, sensory-grounded, knowledge-enriched, and human-based semantic representations. Our analysis includes both mesoscale and global scale of studying graph structures. Since classifying the semantic representations based on their graph structures may be confounded by the graph sizes, we introduced a novel statistical approach that improved the clustering of semantic representations while considering the effect of graph size in the comparisons. Using this framework, we found that in human-based semantic graphs, most central nodes are the most frequent words in English, while this is not the case for representations built from distributional semantic models. Moreover, comparing base representations to their combined counterparts we found that adding extra knowledge to a base representation can result in various structural changes. For instance, adding visual semantic knowledge to a distributional space can decrease the probability of forming semantic groups, however, adding human-based knowledge can increase this probability. Finally, we observed that applying the suggested statistical comparison framework yields better clustering for different-sized semantic graphs. To the best of our knowledge, this is the first study aimed to compare semantic representations in such a comprehensive manner in which we included 7 different models of semantic representations. Our findings can have multiple implications for developing more effective and interpretable models in natural language processing (NLP) and for understanding how to combine the strengths of different representations to improve performance on a wide range of NLP tasks. Moreover, the statistical method we introduced in this study in order to compare different semantic graphs that are varying in their size is a pioneering effort that can be used as a general method of comparison in network science. In this regard, some suggestions have been made on how the findings of this work can contribute to future studies.

Keywords Semantic Representation, Pre-trained Language Models, Graph Theory, Complex Network, Different-Sized Networks Comparison, Configuration Model

University of Tehran
Faculty of New Sciences and Technologies

Network Science-based Semantic Representation Analysis

A Thesis submitted to the Graduate Studies Office
In partial fulfillment of the requirements for
The degree of Master of Science
in Computational Linguistics

By:

Mohanna Hoveyda

Supervisors:

Dr. Mostafa Salehi and Dr. Mahmood Bijankhan

Advisor:

Dr. Paulino Villas Boas

January 2022