



دانشگاه صنعتی شریف  
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی  
مهندسی کامپیوتر

# تشخیص اهمیت اخبار فارسی با استفاده از مدل‌های زبانی بزرگ

نگارش

شایان صالحی

استاد راهنما

دکتر مهدی جعفری

بهمن ۱۴۰۴

## سپاس

از استاد بزرگوارم، دکتر جعفری به خاطر زحمات و راهنمایی‌هایی که در طول این پروژه داشته‌اند  
متشکرم و همچنین از دانشجوی دکترا ایشان، آقای معین سلیمی به خاطر زمان و راهنمایی‌هایی که برای  
پیش‌بردن پروژه انجام داده‌اند قدردانم.

## چکیده

این پروژه به بررسی قدرت تشخیص اهمیت یک خبر فارسی توسط مدل‌های زبانی بزرگ پرداخته و قدرت یادگیری از محتوا، قدرت استدلال و قدرت تفکر آن را ارزیابی کرده است. در ابتدا، از دادگان علائم‌گذاری‌شده توسط افراد در حوزه‌های مختلف از جمله ورزشی، سیاسی، اجتماعی، پزشکی و فرهنگی استفاده و محیطی برای ارزیابی مدل‌های زبانی بزرگ توسعه داده شده است. در این محیط مدل‌های مختلف موجود بررسی و ارزیابی شده و در نهایت با تمام حالات مختلف و شرایط مختلف، قدرت تحلیل آنها در زبان فارسی و انگلیسی بررسی شده است. این پروژه نشان‌دهنده که دستورهای<sup>۱</sup> شامل زنجیره تفکر<sup>۲</sup> و درخت تفکر<sup>۳</sup> باعث بهبود کارایی مدل‌ها و همچنین روش تنظیم نمادها<sup>۴</sup> باعث حساسیت بسیار زیاد به پرسش داده شده و محتوای آن می‌شود.

**کلیدواژه‌ها:** مدل‌های زبانی بزرگ، پردازش زبان‌های طبیعی، یادگیری ماشین، تشخیص اهمیت اخبار

---

<sup>1</sup>Prompt

<sup>2</sup>Chain-of-Thoughts

<sup>3</sup>Tree-of-Thoughts

<sup>4</sup>Symbol Tuning

# فهرست مطالب

۱	مقدمه	۱
۱-۱	تعریف مسئله چگونگی بررسی مهم بودن یک خبر . . . . .	۱
۲-۱	اهمیت موضوع تشخیص اهمیت اخبار . . . . .	۲
۳-۱	ادبیات به کار رفته در این پژوهش . . . . .	۲
۴-۱	اهداف پژوهش . . . . .	۲
۲	مفاهیم اولیه	۴
۱-۲	مدل‌های زبانی بزرگ . . . . .	۴
۲-۲	یادگیری درونی . . . . .	۴
۳-۲	تنظیم بر اساس دستورالعمل . . . . .	۵
۴-۲	درخواست‌های سامانه و کاربر . . . . .	۵
۵-۲	مهندسی درخواست . . . . .	۶
۳	کارهای پیشین	۷
۱-۳	تشخیص اهمیت اخبار . . . . .	۷
۱-۱-۳	رویکردهای کلاسیک . . . . .	۷
۲-۱-۳	رویکردهای تشخیص اهمیت با استفاده از یادگیری ماشین . . . . .	۸
۳-۱-۳	رویکرد یادگیری عمیق . . . . .	۸
۴-۱-۳	استفاده از مدل‌های زبانی بزرگ . . . . .	۸

۲-۳	تنظیم بر اساس دستورالعمل و تنظیم نمادین	۹
۳-۳	یادگیری چندنمونه‌ای در تشخیص اهمیت اخبار	۹
۴-۳	تشخیص اهمیت اخبار فارسی	۹
۵-۳	سامانه‌های درخواست وابسته به پرسش	۱۰
۴	روش پیشنهادی	۱۱
۱-۴	انواع دستورالعمل‌های توسعه‌داده شده	۱۱
۱-۱-۴	دستورالعمل‌های وانیلا	۱۱
۵	نتایج جدید	۱۲
۶	نتیجه‌گیری	۱۳
	مراجع	۱۴
	واژه‌نامه	۱۷
آ	مطالب تکمیلی	۱۸
آ-۱	دستورالعمل‌های به کار گرفته شده	۱۸

# فهرست جداول

# فهرست تصاویر

# فصل ۱

## مقدمه

در دنیای رو به پیش رفت روزمره، حجم عظیمی از اخبار شبانه‌روز به سمت کاربران روانده می‌شود. در این حین می‌دانیم که بسیاری از این اخبار مبنای درستی نداشته و بسیاری نیز برای کاربران بسیار اهمیت کمی دارد. با معرفی یک بستر که بتوان به وسیله آن اخبار مهم به خصوص با توجه به فرهنگ ایرانیان تشخیص داده خود یک چالش بزرگ اما بسیار کاربردی است. در اینجا با استفاده و بهره‌گیری از مدل‌های زبانی بزرگ و دانش که توسط آنها جمع‌آوری شده است به انجام این امر پرداختیم. در ادامه همچنین چالش‌های این مدل‌ها و منطبق نبودن آن طبق فرهنگ و عادات ایرانیان بررسی می‌کنیم و با ارائه روش یادگیری چند نمونه<sup>۱</sup>، این مشکل را برای طرف می‌کنیم.

## ۱-۱ تعریف مسئله چگونگی بررسی مهم بودن یک خبر

مسئله به این شکل تعریف می‌شود که یک خبر در هر دسته‌ای که قرار داشته باشد یا دارای اهمیت بالا یا برچسب ۱ و یا دارای اهمیت پایین و برچسب ۰ است. با دادگان جمع‌آوری شده و برچسب‌گذاری‌های انسانی روی آنها، به ۵۵۰۹ داده آموزش و ۱۱۸۰ داده تست و ارزیابی رسیده، که با استفاده از آنها مدل‌ها توصیه نمونه براساس شباهت تعریف شده است و هدف آن است که مدل بتواند اهمیت خبر (۰ یا ۱) را تشخیص دهد و به کاربر اعلام کند.

---

<sup>1</sup>Few-Shot Learning



## ۲-۱ اهمیت موضوع تشخیص اهمیت اخبار

از اهمیت این کار و محیط توسعه داده شده می توان به موارد زیر اشاره کرد:

- تسهیل پیگیری اخبار برای کاربران، از آنجایی که این محیط توان تشخیص اخبار مهم در دسته ها مختلف را داشته، می توان برای کاربران صرفا اخبار مهم را دسته بندی کرده و آنها با خواندن این اخبار در وقت خود نسبت به خواندن مطالب بی اهمیت صرفه جویی خواهند کرد.
- بررسی قدرت استدلال و تفکر مدل های زبانی بزرگ، از آنجایی تشخیص اهمیت یک خبر کار نسبتا پیچیده ای برای این مدل ها احتساب می شود، این بستر فراهم شده است که قدرت استدلال و تحلیل مدل های مختلف در شرایط گوناگون ارزیابی و اعلام شود.
- در این کار، روش هایی برای بهبود و بهینه کردن دقت این مدل ها پیشنهاد و بررسی شده که در جنبه های دیگری غیر از تشخیص اخبار مهم می توان کمک کننده باشد و به کار گرفته شود. از جمله اینها مسئله طبقه بندی و یادگیری محتوای دستور یا درخواست داده شده به مدل های زبانی بزرگ است.

## ۳-۱ ادبیات به کار رفته در این پژوهش

از آنجایی که بخش هایی از این پروژه الهام گرفته و ادامه کار تنظیم نمادها [۱] بوده از ادبیات این کار نیز در اینجا استفاده شده است. تنظیم نمادها عبارت است از روشی که به جای برچسب های اصلی که در اینجا همان ۰ یا ۱ هستند، یک رشته از نمادها همانند !، #، & و کاراکترهای دیگر جایگزین شود و مدل نتواند به دانش پیشینه خود اتکا کند.

## ۴-۱ اهداف پژوهش

اهداف این پژوهش صورت گرفته به دو قسمت کلی تقسیم می شود:

- ابتدا با ساختار و تعریف اخبار مهم پرداخته شده است، در این پژوهش نتیجه ها و بررسی های انجام شده حاکی این موضوع است که اخبار در دسته های گوناگون و برای اشخاص با فرهنگ های مختلف اهمیت متفاوتی دارد. بنابراین انجام یک مسئله طبقه بندی روی آنها کار آسانی نبوده و با بهبودهای

انجام شده در این پژوهش، مسیری برای پژوهش‌های بعدی در جهت رسیدن که دقت بالا با در نظر گرفتن تمام این شرایط فراهم کند.

- مدل‌های زبان بزرگ که کانون اصلی توجه این پژوهش بوده است در این مسئله خاص به طور کامل بررسی شده و تمامی نقاط ضعف و قوت این مدل‌ها در تشخیص اهمیت اخبار بررسی شده است. همچنین تفاوت قدرت استدلال این مدل‌ها در زبان فارسی با انگلیسی مورد مقایسه قرار گرفته که خود می‌تواند مورد استناد برای پژوهش‌های آینده در این زمینه قرار گیرد.

## فصل ۲

# مفاهیم اولیه

در اینجا به مفاهیم اصلی به کار برده شده در این پروژه، و بررسی کاربرد و پیشینه آن می‌پردازیم.

## ۱-۲ مدل‌های زبانی بزرگ

مدل‌های زبانی بزرگ<sup>۱</sup> به سامانه‌های هوش مصنوعی گفته می‌شوند که بر اساس پردازش زبان طبیعی طراحی شده‌اند و قادر به تولید و درک متن‌های انسانی در مقیاس وسیع هستند. این مدل‌ها با استفاده از حجم بسیار زیادی از داده‌های متنی آموزش می‌بینند و می‌توانند وظایف متنوع زبانی، از جمله ترجمه، خلاصه‌سازی و پاسخ به پرسش‌ها را انجام دهند.

مفهوم مدل‌های زبانی از دهه ۱۹۸۰ با ظهور الگوریتم‌های احتمالاتی ساده آغاز شد. با معرفی شبکه‌های عصبی در دهه ۱۹۹۰ و توسعه یادگیری عمیق در دهه ۲۰۱۰، مدل‌هایی مانند ترانسفورمرها و سیستم‌هایی نظیر جی‌پی‌تی (نسل اول تا سوم) و مدل‌های مشابه توانستند به کارایی فوق‌العاده‌ای دست یابند. [۲] افزایش قدرت محاسباتی و دسترسی به داده‌های بیشتر، این پیشرفت‌ها را تسهیل کرد.

## ۲-۲ یادگیری درونی

یادگیری درون‌متنی<sup>۲</sup> به توانایی یک مدل زبانی اشاره دارد که بتواند بر اساس نمونه‌هایی که در همان متن ورودی ارائه می‌شود، وظایف جدیدی را یاد بگیرد. در این روش، نیاز به آموزش دوباره مدل وجود ندارد،

<sup>۱</sup>Large Language Models

<sup>۲</sup>In-Context Learning

بلکه مدل از اطلاعات داده شده در همان لحظه استفاده می‌کند. [۳]

این مفهوم در اوایل دهه ۲۰۲۰ با توسعه مدل‌هایی مانند جی‌پی‌تی ۳ به وضوح مطرح شد. این مدل‌ها نشان دادند که بدون نیاز به آموزش دوباره، می‌توانند تنها با ارائه نمونه‌هایی در ورودی، وظایف مختلفی را انجام دهند. این پیشرفت‌ها نقطه عطفی در ساده‌سازی استفاده از مدل‌های زبانی محسوب می‌شوند.

## ۳-۲ تنظیم بر اساس دستورالعمل

تنظیم بر اساس دستورالعمل<sup>۳</sup> فرآیندی است که در آن یک مدل هوش مصنوعی با استفاده از داده‌هایی آموزش می‌بیند که حاوی دستورالعمل‌های خاصی برای انجام وظایف مختلف هستند. [۴] هدف این روش بهبود عملکرد مدل در درک و اجرای دستورالعمل‌هاست.

ایده این روش از مفاهیم یادگیری انتقالی نشأت گرفته است. در سال‌های اخیر، با توجه به توانایی مدل‌های بزرگ زبانی در تعمیم وظایف، محققان تلاش کردند تا این مدل‌ها را با داده‌های حاوی دستورالعمل بهبود دهند. پروژه‌هایی مانند اجرای دستور عمل در مدل‌های جی‌پی‌تی [۵] نشان‌دهنده موفقیت این رویکرد هستند.

## ۴-۲ درخواست‌های سامانه و کاربر

درخواست‌های سامانه<sup>۴</sup> و کاربر به متونی اطلاق می‌شود که برای هدایت مدل زبانی به سمت تولید پاسخ مناسب استفاده می‌شوند. درخواست سامانه معمولاً وظیفه مشخص کردن قواعد کلی را دارد، در حالی که درخواست کاربر هدف یا سؤال خاصی را بیان می‌کند. [۶]

این مفهوم با گسترش استفاده از مدل‌های زبانی در تعاملات انسانی به وجود آمد. اولین تلاش‌ها برای تعریف و تمایز این دو نوع درخواست در توسعه رابط‌های کاربری تعاملی و چت‌بات‌ها مشاهده شد. این ایده در مدل‌های زبانی بزرگ تکامل یافت.

---

<sup>3</sup>Instruction Tuning

<sup>4</sup>System Prompt

## ۵-۲ مهندسی درخواست

مهندسی درخواست<sup>۵</sup> به هنر و دانش طراحی درخواست‌ها برای هدایت مدل‌های زبانی جهت تولید پاسخ‌های دقیق و مفید اشاره دارد. این فرآیند شامل ایجاد ورودی‌هایی است که بتوانند بهترین نتیجه ممکن را از مدل دریافت کنند.

این مفهوم با ظهور مدل‌های زبانی پیچیده و نیاز به بهره‌برداری بهتر از توانایی‌های آن‌ها مطرح شد. در سال‌های اخیر، مقالات و ابزارهای بسیاری برای استانداردسازی و بهبود این فرآیند ارائه شده است. [۷] مهندسی درخواست در زمینه‌های مختلف، از پژوهش گرفته تا صنعت، نقش کلیدی ایفا می‌کند.

---

<sup>۵</sup>Prompt Engineering

## فصل ۳

### کارهای پیشین

همواره در طول زمان بررسی اهمیت اخبار چه در زبان فارسی و چه در زبان انگلیسی یک دغدغه و یک کار مبهم بوده است. از آنجایی که اهمیت یک خبر وابسته به عوامل مختلف همانند فرهنگ، موقعیت جغرافیایی، سلائق شخصی و دیدگاه‌های کاربران بوده در نگاه اول به نظر این کار، ناممکن می‌رسد. اما پژوهش‌های اخیر نشان داده است که با استفاده از دادگان‌های برچسب‌گذاری شده و استفاده از یادگیری چند نمونه‌ای می‌توان به نتایج قابل قبولی برای این قسمت رسید.

در اینجا به روش‌های مختلف که در گذشته برای بررسی اهمیت اخبار توسعه داده شده است پرداخته شده است و سپس مسیرهای مختلف بررسی و آنالیز این طبقه‌بندی را در مدل‌های زبانی بزرگ بیان شده است.

### ۳-۱ تشخیص اهمیت اخبار

این بخش به بررسی روش‌های مختلفی می‌پردازد که در طول زمان برای تشخیص اهمیت اخبار استفاده شده‌اند. این روش‌ها شامل رویکردهای کلاسیک، یادگیری ماشین و هوش مصنوعی، یادگیری عمیق و در نهایت مدل‌های زبانی بزرگ هستند.

### ۳-۱-۱ رویکردهای کلاسیک

در روش‌های کلاسیک، تشخیص اهمیت اخبار بیشتر بر اساس معیارهای دستی انجام می‌شد. از معیارهایی مانند طول خبر، تعداد دفعات ذکر شدن یک موضوع در منابع مختلف، یا تحلیل‌های آماری ساده برای این

کار استفاده می‌شد. [۸] این روش‌ها به دلیل محدودیت در قابلیت درک معنایی متون، کارایی پایینی در مسائل پیچیده داشتند.

### ۳-۱-۲ رویکردهای تشخیص اهمیت با استفاده از یادگیری ماشین

با ظهور الگوریتم‌های یادگیری ماشین<sup>۱</sup>، از مدل‌هایی مانند ماشین بردار پشتیبان<sup>۲</sup>، دسته‌بند بیزین ساده و جنگل‌های تصادفی<sup>۳</sup> [۹] برای تحلیل اخبار و تشخیص اهمیت آن‌ها استفاده شد. [۱۰] این روش‌ها از ویژگی‌های استخراج‌شده مانند تعداد کلمات کلیدی، میزان تعامل خبرها و تعداد نقل‌قول‌ها بهره می‌بردند.

### ۳-۱-۳ رویکرد یادگیری عمیق

با توسعه یادگیری عمیق<sup>۴</sup>، استفاده از شبکه‌های عصبی مانند شبکه‌های بازگشتی و شبکه‌های توجه‌محور برای درک معنایی متون و تحلیل اخبار رواج یافت. مدل‌هایی نظیر LSTM و GRU توانستند با درک وابستگی‌های طولانی‌مدت در متن [۱۱]، عملکرد چشمگیری ارائه دهند.

### ۳-۱-۴ استفاده از مدل‌های زبانی بزرگ

مدل‌های زبانی بزرگ با پیش‌آموزش بر داده‌های گسترده، توانایی تحلیل متون خبری را با دقت بالا فراهم کرده‌اند. [۲] این مدل‌ها با توجه به پیکربندی و حجم داده‌های آموزشی، قادرند وظایف مختلف را به صورت چندمنظوره انجام دهند.

در کارهای پیشین انجام شده بررسی شده که رفتار مدل‌های زبانی بزرگ در تشخیص اخبار جعلی چگونه بوده است. به خصوص در پژوهش‌های قبلی [۱۲] با تعریف بازیگر خوب و بد، سعی بر ارزیابی این نوع اخبار داشته و نشان می‌دهد که این مدل‌ها توانایی مناسب جهت تشخیص اخبار جعلی در شرایط از پیش تعریف شده مناسب خواهند داشت.

همچنین کارهای فراتری نسبت به صرفاً اتکا کردن به پردازش متن انجام شده است، به طوری که با بهره‌گیری همزمان از مدل‌های تصویری مانند CLIP اهمیت اخبار براساس محتوای تصویری، ویدیوی و صوتی به همراه متن آنها نیز بررسی شود. [۱۳]

<sup>1</sup>Machine Learning

<sup>2</sup>SVM

<sup>3</sup>Random Forest Tree

<sup>4</sup>Deep Learning

## ۲-۳ تنظیم بر اساس دستورالعمل و تنظیم نمادین

تنظیم بر اساس دستورالعمل به آموزش مدل‌های زبانی بزرگ با داده‌هایی اشاره دارد که شامل دستورالعمل‌های دقیق برای انجام وظایف هستند. [۱۴] این روش باعث می‌شود مدل‌ها بتوانند وظایف مشخصی مانند دسته‌بندی اهمیت اخبار را با دقت بیشتری انجام دهند. از سوی دیگر، تنظیم نمادین<sup>۵</sup> شامل استفاده از اطلاعات ساختاریافته مانند نمودارهای دانش یا نمایش‌های معنایی برای تقویت عملکرد مدل‌ها است.

## ۳-۳ یادگیری چندنمونه‌ای در تشخیص اهمیت اخبار

در این روش، مدل‌ها با تعداد بسیار کمی از نمونه‌های آموزشی، وظایف خود را یاد می‌گیرند. این ویژگی در تحلیل اخبار و تشخیص اهمیت آن‌ها، به‌ویژه در مواقعی که داده‌های آموزشی محدود است، کاربرد دارد. مدل‌هایی مانند Aya توانسته‌اند نشان دهند که تنها با چند نمونه ورودی می‌توانند وظایف پیچیده را انجام دهند. [۱۵]

## ۴-۳ تشخیص اهمیت اخبار فارسی

در گذشته، تلاش‌هایی برای توسعه مدل‌های تشخیص اهمیت اخبار فارسی صورت گرفته است. این تلاش‌ها بیشتر بر اساس روش‌های یادگیری ماشین کلاسیک بوده و از ویژگی‌های زبانی خاص فارسی مانند ریشه‌یابی و تحلیل صرفی بهره گرفته‌اند. [۱۶] با این حال، استفاده از مدل‌های زبانی بزرگ برای زبان فارسی هنوز در مراحل ابتدایی قرار دارد.

همچنین این پژوهش، ادامه مسیر کار خبرچین [۱۷] بوده که با استفاده از مدل‌های مبنی بر معماری ترانسفورمر سعی داشته که به بررسی اهمیت اخبار بپردازد. در این پژوهش سعی شده با استفاده از مدل‌های زبانی بزرگ رویکرد کلی‌تری نسبت به بررسی اهمیت اخبار ارائه شود که بتواند بستر جامع‌تری برای طبقه‌بندی این حوزه فراهم کند.

---

<sup>۵</sup>Symbol Tuning



### ۵-۳ سامانه‌های درخواست وابسته به پرسش

این سامانه‌ها با طراحی درخواست‌هایی که وابسته به موضوع پرسش هستند، قادرند نتایج بهینه‌ای در تشخیص اهمیت اخبار ارائه دهند. برای این منظور، از مهندسی درخواست استفاده می‌شود تا مدل‌های زبانی بزرگ بتوانند بر اساس متن ورودی و هدف پرسش، خروجی مطلوبی تولید کنند. [۱۸]

پژوهش‌های اخیر در این زمینه انجام شده است که نشان می‌دهد استفاده از دستورهای مختلف بسته به ورودی کاربر می‌تواند نتایج ثمربخش‌تر به ارقام بیاورد. اگرچه در یکسری پژوهش‌های به این رویکرد یادگیری تقویتی<sup>۶</sup> الحاق می‌شود [۱۹] اما بیشتر یک سیستم در پس‌زمینه بوده که بتواند بهترین دستور را با توجه به ورودی و محتوا تشخیص دهد.

---

<sup>6</sup>Reinforcement Learning

## فصل ۴

### روش پیشنهادی

در این قسمت به روش‌های توسعه داده‌شده و نحوه به دست آمدن نتایج و خروجی‌ها می‌پردازیم. فرآیند به این صورت طی می‌شود که ابتدا دستورالعمل مناسب براساس شرط‌های مشخص شده انتخاب می‌شود و سپس در صورت نیاز نمونه‌های مشابه به خبر مورد نظر در دستور آمده و سپس مدل‌های زبانی بزرگ ۸ یا ۹ میلیارد پارامتر روی کارت گرافیکی بالا آمده و به صورت دسته‌های ۱۰ تایی نتایج از خروجی این مدل‌ها ذخیره شده و تحلیل‌ها روی آن انجام می‌شود.

#### ۴-۱ انواع دستورالعمل‌های توسعه داده شده

در اینجا به چهار حالت پرامپت‌ها یا همان دستورالعمل‌ها را که به کار گرفته شده است پرداخته می‌شود و کاربرد و اهداف هرکدام بررسی می‌شود. یک دستورالعمل  $p_i$  از یک مجموعه  $P = \{p_1, p_2, \dots, p_n\}$  انتخاب می‌شود و براساس آن نتایج خروجی که به صورت برچسب‌های  $l_j = \langle 0, 1 \rangle$  مشخص می‌شود خروجی گرفته می‌شود.

#### ۴-۱-۱ دستورالعمل‌های وانیلا

دستورالعمل‌های وانیلا<sup>۱</sup> که همان دستورالعمل‌های خام بوده صرفاً اطلاعات مورد نیاز را برای مدل‌های زبانی بزرگ تهیه و توضیح می‌دهد.

---

<sup>۱</sup>Vanila

## فصل ۵

### نتایج جدید

در این فصل نتایج جدید به دست آمده در پایان نامه توضیح داده می شود. در صورت نیاز می توان نتایج جدید را در قالب چند فصل ارائه نمود. همچنین در صورت وجود پیاده سازی، بهتر است نتایج پیاده سازی را در فصل مستقلی پس از این فصل قرار داد.

## فصل ۶

### نتیجه‌گیری

در این فصل، ضمن جمع‌بندی نتایج جدید ارائه‌شده در پایان‌نامه یا رساله، مسائل باز باقی‌مانده و همچنین پیشنهادهایی برای ادامه‌ی کار ارائه می‌شوند.

# Bibliography

- [1] J. Wei, L. Hou, A. Lampinen, X. Chen, D. Huang, Y. Tay, X. Chen, Y. Lu, D. Zhou, T. Ma, and Q. V. Le. Symbol tuning improves in-context learning in language models, 2023.
- [2] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.
- [3] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. Language models are few-shot learners, 2020.
- [4] J. Wei, M. Bosma, V. Y. Zhao, K. Guu, A. W. Yu, B. Lester, N. Du, A. M. Dai, and Q. V. Le. Finetuned language models are zero-shot learners, 2022.
- [5] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, and R. Lowe. Training language models to follow instructions with human feedback, 2022.
- [6] S. Gao, A. Sethi, S. Agarwal, T. Chung, and D. Hakkani-Tur. Dialog state tracking: A neural reading comprehension approach, 2019.
- [7] L. Reynolds and K. McDonell. Prompt programming for large language models: Beyond the few-shot paradigm, 2021.
- [8] H. P. Luhn. The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2(2):159–165, 1958.

- [9] M. Felicilda, A. Geriane, V. Agustin, M. C. Blanco, J. Morano, L. Mahusay, and J. Guialil. Enhancement of random forest algorithm applied in fake news detection. *World Journal of Advanced Research and Reviews*, 22:1075–1079, 05 2024.
- [10] Y. Yang and J. O. Pedersen. A comparative study on feature selection in text categorization. In *Proceedings of the Fourteenth International Conference on Machine Learning, ICML '97*, page 412–420, San Francisco, CA, USA, 1997. Morgan Kaufmann Publishers Inc.
- [11] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, Nov. 1997.
- [12] B. Hu, Q. Sheng, J. Cao, Y. Shi, Y. Li, D. Wang, and P. Qi. Bad actor, good advisor: Exploring the role of large language models in fake news detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(20):22105–22113, Mar. 2024.
- [13] J. Wang, Z. Zhu, C. Liu, R. Li, and X. Wu. Llm-enhanced multimodal detection of fake news. *PLOS ONE*, 19, 10 2024.
- [14] V. Sanh, A. Webson, C. Raffel, S. H. Bach, L. Sutawika, Z. Alyafeai, A. Chaffin, A. Stiegler, T. L. Scao, A. Raja, M. Dey, M. S. Bari, C. Xu, U. Thakker, S. S. Sharma, E. Szczechla, T. Kim, G. Chhablani, N. Nayak, D. Datta, J. Chang, M. T.-J. Jiang, H. Wang, M. Manica, S. Shen, Z. X. Yong, H. Pandey, R. Bawden, T. Wang, T. Neeraj, J. Rozen, A. Sharma, A. Santilli, T. Fevry, J. A. Fries, R. Teehan, T. Bers, S. Biderman, L. Gao, T. Wolf, and A. M. Rush. Multitask prompted training enables zero-shot task generalization, 2022.
- [15] P. Zheng, H. Chen, S. Hu, B. Zhu, J. Hu, C.-S. Lin, X. Wu, S. Lyu, G. Huang, and X. Wang. Few-shot learning for misinformation detection based on contrastive models. *Electronics*, 13(4), 2024.
- [16] M. Heydari, M. Khazeni, and M. A. Soltanshahi. Deep learning-based sentiment analysis in persian language. In *2021 7th International Conference on Web Research (ICWR)*, page 287–291. IEEE, May 2021.
- [17] H. H. Hemati, A. Lagzian, M. S. Sartakhti, H. Beigy, and E. Asgari. Khabarchin: Automatic detection of important news in the persian language, 2023.
- [18] L. Reynolds and K. McDonell. Prompt programming for large language models: Beyond the few-shot paradigm, 2021.

- [19] H. Sun, A. Hüyük, and M. van der Schaar. Query-dependent prompt evaluation and optimization with offline inverse rl, 2024.

# واژه‌نامه

## ت

تنظیم براساس دستورالعمل Insturction Tuning....  
تنظیم نمادین Symbol Tuning .....

## ق

قطعی Deterministic.....

## ج

جنگل‌های تصادفی Random Forest Tree.....

## م

مدل‌های زبانی بزرگ Large Language Models....

مهندسی درخواست Prompt Engineering.....

ماشین بردار پشتیبان SVM.....

## د

دستور Prompt .....

درخت تفکر Tree-of-Thoughts.....

درخواست سامانه System Prompt.....

## و

دستورالعمل خام Vanila Prompt.....

## ی

یادگیری چند نمونه Few-Shot Learning.....

یادگیری درون‌متنی In-Context Learning.....

یادگیری ماشین Machine Learning .....

یادگیری عمیق Deep Learning .....

یادگیری تقویتی Reinforcement Learning.....

## ز

زنجیره تفکر Chain-of-Thoughts.....

## ص

صدق‌پذیری Satisfiability.....



# پیوست آ

## مطالب تکمیلی

در اینجا تمام محتوای تکمیلی پژوهش از جمله دستورالعمل‌های استفاده شده در مراحل مختلف برای مدل‌های زبانی بزرگ و نتایج دسته‌بندی شده قرار گرفته شده است.

### آ-۱ دستورالعمل‌های به کار گرفته شده

## **Abstract**

This work examines the capability of large language models (LLMs) to measure the importance of Persian news, evaluating their learning ability from content, reasoning skills, and overall cognitive capacities. Initially, annotated datasets were collected from various domains, including sports, politics, social issues, medicine, and culture, to develop an evaluation framework for LLMs. Within this framework, various existing models were analyzed and assessed under different scenarios and conditions to evaluate their analytical performance in both Persian and English. The findings indicate that prompts incorporating Chain-of-Thoughts and Tree-of-Thoughts significantly improve the models' performance. Additionally, the Symbol Tuning method enhances sensitivity to the input queries and their content.

**Keywords:** Large Language Models, Natural Language Processing, Machine Learning, News Importance Detection



Sharif University of Technology  
Department of Computer Engineering

B.Sc. Thesis

# **News Importance Detection With the Use of Large Language Models**

By:

**Shayan Salehi**

Supervisor:

**Dr. Mahdi Jafari**

January 2025