تحلیل احساسات مبتنی بر جنبه با استفاده از شبکهٔ رمزگذار توجه

سمیه کریمی * و فاطمه جعفری نژاد
دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود، ایران

چکیده

پردازش زبان طبیعی به‌طور قابل توجهی در حال رشد و با ظهور تارنمای جهانی و موتورهای جستجو بسیار مورد توجه قرار گرفته است و پژوهش‌گران شاهد انفجاری در اطلاعات به زبان‌های مختلف شدند. تحلیل احساسات یکی از فعال‌ترین زمینه‌های مطالعاتی در پردازش زبان طبیعی است که بر طبقه بندی متن تمرکز دارد و به‌منظور شناسایی، استخراج و تجزیه و تحلیل اطلاعات ذهنی از منابع متنی استفاده می‌شود. تحلیل احساسات مبتنی بر جنبه یک روش تحلیل متن است که نظرات را بر اساس جنبه طبقه‌بندی و احساسات مربوط به هر جنبه را مشخص می‌کند. این تحلیل می‌تواند برای تحلیل خودکار بازخورد نظرات مشتریان به بخش‌های مختلف کالا یا خدمات مورد استفاده قرار گیرد و به کارفرمایان برای تمرکز بر نقاط نیازمند ارتقای کیفیت کمک کند. در این مقاله به معرفی یک معماری جدید مبتنی بر یادگیری عمیق برای تحلیل احساسات مبتنی بر جنبه خواهیم پرداخت. این معماری از یک مدل از دولایهٔ رمزگذار توجه (که یک جایگزین قابل موازی‌سازی و تعاملی LSTM است و برای محاسبه حالت‌های پنهان جاسازی‌های ورودی اعمال می‌شود) استفاده خواهد کرد. آزمایش این معماری روی سه مجموعه‌دادهٔ مختلف شامل رستوران ها و لپتاپ‌ها SemEval 2014 Task 4 و مجموعه‌داده ACL 14 Twitter است که در هر سه مجموعه‌داده، قطبیت احساسات مثبت، خنثی و منفی است، انجام شده‌است که مقایسه آن با روش‌های مدرن تحلیل احساس مبتنی بر جنبه، دقت بالای این روش را نشان خواهد داد. برای نمونه، تحلیل احساس مبتنی بر جنبه روی مجموعه داده لپتاپ، ۷۹.۱۵ درصد دقت را نشان داده که نسبت به روش‌های مدرن ۴.۲۴ درصد دقت را بالا برده است.

واژگان کلیدی: تجزیه و تحلیل احساس مبتنی بر جنبه، یادگیری عمیق، شبکه رمزگذار توجه، توجه چندسر

Aspect-Based Sentiment Analysis using the Attentional Encoder Network

Somayeh Karimi* And Fatemeh Jafarinejad
Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran

Abstract

Natural language processing is growing significantly and has gained much attention with the advent of the World Wide Web and search engines, and researchers have witnessed an explosion of information in different languages. Sentiment analysis is one of the most active fields of study in natural language processing that focuses on text classification and is used to identify, extract and analyze subjective information from text sources. Aspect-based sentiment analysis is a text analysis technique that classifies comments by aspect and identifies the sentiment associated with each aspect. This analysis can be used to automatically analyze the feedback of customers' comments to different parts of goods or services and help employers to focus on points that need quality improvement. In this paper, we will introduce a new architecture based on deep learning for aspect-based sentiment analysis. This architecture will use an attention-encoder network-based model with multiple multi-head attention and a pointwise convolutional transform (which is a parallelizable and interactive alternative to LSTM and is applied to compute hidden states of input embeddings). Testing this architecture on three different datasets, including restaurants and laptops, SemEval 2014 Task 4 and ACL 14 Twitter dataset, in all three

* نویسندهٔ عهده‌دار مکاتبات
* Corresponding author

• تاریخ ارسال مقاله: ۱۴۰۱/۶/۱۰ • تاریخ پذیرش: ۱۴۰۲/۹/۲۰ • تاریخ انتشار: ۱۴۰۲/۱۲/۲۹ • نوع مطالعه: بنیادی

[DOI: 10.61186/jsdp.20.4.121]

datasets, the polarity of emotions is positive, neutral and negative, which is compared with modern methods of sentiment analysis. Based on the aspect, it will show the high accuracy of this method. For example, the aspect-based sentiment analysis on the Laptop dataset has shown 79.15% accuracy, which has increased the accuracy by 4.24% compared to modern methods.

Keywords: Aspect-Based Sentiment Analysis (ABSA)، Deep Learning, Attentional Encoder Network، Multi-Head Attention (MHA)

۱- مقدمه

در سال‌های اخیر پژوهش‌های گسترده‌ای در زمینهٔ تحلیل احساسات انجام شده‌است. طبق آمار WoS، در سال‌های اخیر روند روبه‌رشد ارجاعات به مقالات تحلیل احساسات رشد قابل توجهی داشته که نشان‌دهندهٔ اهمیت چشم‌گیر این موضوع است[۱]. تجزیه و تحلیل احساسات[1] گاهی به‌عنوان نظرکاوی یا عقیده‌کاوی یا هوش مصنوعی احساسات شناخته می‌شود که با استفاده از روش‌های پردازش زبان طبیعی، تجزیه و تحلیل متن، زبان‌شناسی محاسباتی، استخراج و سنجش کمیت، برای سنجش میزان مثبت یا منفی‌بودن نظرات کاربران نسبت به یک محصول یا خدمت اشاره دارد. تحلیل احساسات نظرات کاربران اغلب برای اهدافی از بازاریابی گرفته تا پشتیبانی مشتری استفاده می‌شود[۲]. مانند بررسی امکانات و خدمات یک هتل و یا نقد و بررسی یک فیلم و...

اغلب مطالعه تجزیه و تحلیل احساسات در سه سطح جزیی‌تری انجام می‌شود: سطح سند، سطح جمله و سطح جنبه. در سطح سند، یک سند به‌عنوان بیان‌کنندهٔ یک نظر کلی مثبت یا منفی طبقه‌بندی می‌شود در این سطح کل سند را به‌عنوان واحد اطلاعات بنیادی در نظر گرفته می‌شود. به این معنا که هر سند نظرات خود را در مورد یک موجودیت واحد بیان می‌کند[۳].

طبقه‌بندی سطح جمله هر جمله را به‌عنوان یک واحد جداگانه در نظر می‌گیرد و فرض می‌کند که جمله باید فقط یک نظر داشته باشد که در آن یک جمله به‌عنوان خنثی، مثبت یا منفی طبقه‌بندی می‌شود[۴].

تحلیل احساسات سطح جنبه، که به‌عنوان تحلیل احساسات مبتنی بر جنبه نیز شناخته می‌شود، زیرمجموعه‌ای از تجزیه و تحلیل احساسات است که نسبت به دو سطح قبل به‌صورت جزئی‌تری عمل می‌کند. درواقع این روش به نظرات کاربران به‌صورت جزئی‌تر نگاه می‌کند و برای هر یک از موجودیت‌ها یا جنبه‌های مختلف موجود در اظهار نظر، به‌صورت تفکیک‌شده به سنجش قطبیت نظر (مثبت، منفی، یا خنثی‌بودن) آن می‌پردازد[۵].

از آن‌جایی که احساسات در یک جمله می‌تواند پیچیده باشد و در مورد جنبه‌های مختلف، احساسات متفاوتی وجود داشته باشد، تحلیل احساسات مبتنی بر جنبه (ABSA) پیشنهاد شده‌است[۵]؛ لذا این روش روشی مناسب را در اختیار سازمان‌ها و شرکت‌های مختلف قرار می‌دهد تا بازخورد نظرات مشتریان خود را (که محصول یا خدمات آنان را دریافت کرده‌اند) نسبت به جنبه‌های مختلف مورد سنجش و ارزیابی قرار دهند. این نظرات می‌تواند یک جمله یا کامنت در سایت و یا شبکه‌های اجتماعی باشد. ویژگی‌های پیشرفته تحلیل احساسات مبتنی بر جنبه به این کارفرمایان کمک می‌کند تا اطلاعات را از داده‌ها استخراج کنند و بدین ترتیب دیدی جامع از احساسات مشتری نسبت به جنبه‌های مختلف محصول/خدمات و برند خود به‌دست آورند. برای مثال درنظرسنجی مربوط به لپتاپ ، جملهٔ " وزن آن سبک است اما باتری ضعیفی دارد" نسبت به جنبه‌های "وزن" و "باتری" به ترتیب نظرات مثبت و منفی را بیان می کند.

در سال‌های اخیر استفاده از روش‌های یادگیری عمیق موجب افزایش چشم‌گیر دقت تسک‌های پردازش زبان طبیعی شده‌است. در این مقاله نیز به معرفی یک معماری جدید مبتنی بر یادگیری عمیق برای تحلیل احساسات مبتنی بر جنبه خواهیم پرداخت. این معماری از یک مدل مبتنی بر شبکه رمزگذار توجه[2] با چندین توجه چند سر[3] و تبدیل کانولوشن نقطه‌ای[4] (که یک جایگزین قابل موازی‌سازی و تعاملی LSTM[5] است و برای محاسبه حالت های پنهان تعبیه های ورودی اعمال می‌شود) استفاده خواهد کرد. در این معماری ما از روش تعبیه کلمات BERT[۵] استفاده می‌کنیم. مزیت این روش نسبت به روشهای قدیمی تر تعبیه کلمات مانند Word2Vec[۵] و GloVe[۵] در این است که آزمایش معماری نهایی روی سه مجموعه‌دادگان مختلف و مقایسه آن با روش‌های مدرن تحلیل احساس مبتنی بر جنبه، دقت بالای این روش را نشان خواهد داد. به‌عنوان نمونه، تحلیل

[1] sentiment analysis (SA)

[2] Attentional Encoder Network (AEN)
[3] Multi-Head Attention (MHA)
[4] Point-wise Convolution Transformation (PCT)
[5] Long short-term memory (LSTM)

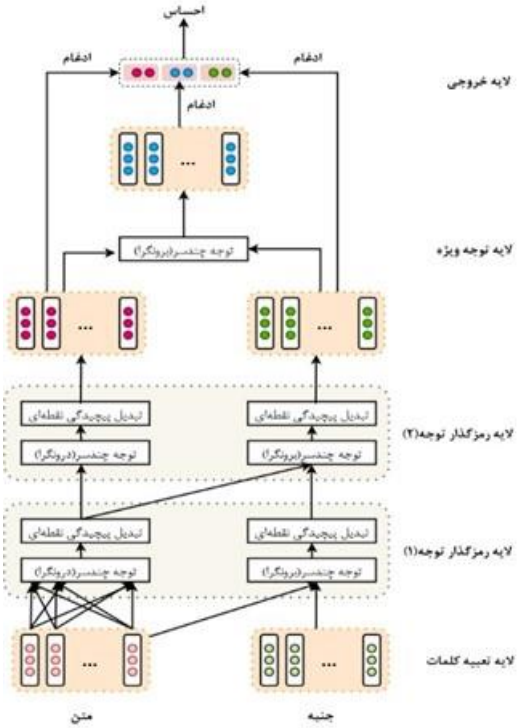
[DOI: 10.61186/jsdp.20.4.121]

احساس مبتنی بر جنبه روی مجموعه‌دادگان رستوران ۸۰.۵۴ درصد دقت را نشان داده است.

در ادامه ساختار مقاله به این صورت است که در بخش ۲ کارهای مرتبط در حوزهٔ تحلیل احساس و تحلیل احساس مبتنی بر جنبه را به اجمال مرور خواهیم کرد. در بخش ۳ معماری پیشنهادی شرح داده خواهد شد. در بخش ۴ به ارزیابی پیشنهادی و ارائه نتایج آن در مقایسه با سایر روش ها می پردازیم و درنهایت در بخش ۵ به نتیجه‌گیری کلی خواهیم پرداخت.

۲-کارهای مرتبط پیشین

در سال‌های اخیر، مدل‌های یادگیری عمیق پیشرفت‌های چشم‌گیری در خودکارسازی تحلیل‌ها ایجاد کرده‌است[۶].



(شکل-۱): معماری کلی مدل پیشنهادی
(Figure- 1) General architecture of the proposed model

با رشد کاربران در وب و شبکه‌های اجتماعی، مردم روزانه ایده‌ها و نظرات خود را در قالب متن، تصویر، فیلم و گفتار به اشتراک می‌گذارند. طبقه‌بندی متن موضوع مهمی است زیرا این متون عظیم از منابع و افراد مختلف با طرز تفکرهای متفاوت تولید می‌شود. پردازش زبان طبیعی و یادگیری عمیق بینش‌های معناداری از احساس یا درک مشتری[1] را مورد بررسی قرار می‌دهد [۷]. هدف از یادگیری عمیق برای تجزیه و تحلیل احساس مشتری این است که روش‌های آن روی حجم زیادی از داده‌های نظرات مشتریان استفاده شود. بسیاری از مدل‌ها، شبکه‌های عصبی کانولوشن[2]، شبکه‌های عصبی بازگشتی[3]، موارد خاص حافظه کوتاه‌مدت برای این منظور قابل استفاده هستند[۸].

دکتر صدر و همکارانشان[۹] یک روش بر پایه ترکیب شبکه عصبی کانولوشن و شبکه عصبی بازگشتی برای تجزیه و تحلیل احساسات موجود در متن را ارائه داده‌اند که در این مدل شبکه عصبی بازگشتی به‌عنوان جایگزین لایه ادغام استفاده شده‌است. هدف از ترکیب و استفاده هم زمان این دو شبکه، ایجاد هم افزایی و بهره‌گیری از مزایای هردوی آنها بوده است.

در ادبیات، طبقه‌بندی احساسات در سطح جنبه به‌طور معمول به‌عنوان چالش طبقه‌بندی نامیده می‌شود. همان‌طور که درقبل گفته شد، دسته‌بندی احساسات در سطح جنبه یک کار طبقه‌بندی دقیق است[۱۰]. Tao Chen و همکاران[۱۱] در این زمینه مطالعاتی داشته‌اند که در آن با معرفی مدل BiLSTM_CRF[4] عبارات هدف را با روش IOB[5] از ورودی استخراج کردند و هر جمله را بر اساس تعداد عبارات هدف به‌صراحت در آن بیان می‌کند، طبقه‌بندی می‌کند؛ سپس، مدل طبقه‌بندی احساسات 1d-CNN را توصیف کرده‌اند که قطبیت احساسات را برای جملات بدون هدف، جملات یک‌هدفه و جملات چندهدفه به‌طور جداگانه پیش‌بینی می‌کند. این مدل یکی از مدل‌های توالی عصبی عمیق است که در آن یک لایه حافظه کوتاه‌مدت دوطرفه (BiLSTM) و یک‌لایه میدان‌های تصادفی شرطی (CRF) ، برای یادگیری دنباله روی هم چیده می‌شوند. BiLSTM یک لایه حافظه کوتاه‌مدت LSTM روبه‌جلو و یک لایه LSTM رو به عقب را به‌منظور یادگیری اطلاعات ترکیب می‌کند.

LSTM در انواع برنامه های NLP عملکرد تحسین‌برانگیزی داشته است. در طبقه‌بندی احساسات وابسته به هدف، TD-LSTMو TC-LSTM[12] با درنظرگرفتن اطلاعات هدف، عملکردی پیشرفته به‌دست آوردند. برخلاف اثربخشی این روش‌ها، تمایز قائل‌شدن قطب‌های احساسات مختلف در سطح جنبه‌های ریز هنوز چالش‌برانگیز است[۱۰].

۳- روش پیشنهادی

در این مقاله به معرفی یک معماری جدید مبتنی بر یادگیری عمیق برای تحلیل احساسات مبتنی بر جنبه

[1] Customer Perception(CP)

[2] Convolutional Neural Network (CNN)
[3] Recurrent Neural Networks(RNN)
[4] Bidirectional Long Short Term Memory with a Conditional Random Feild
[5] Inside–outside–beginning (tagging)

[DOI: 10.61186/jsdp.20.4.121]

می‌پردازیم. شکل (۱) به‌صورت نمادین این معماری را نشان می‌دهد. طبق شکل، مدل پیشنهادی از چهار قسمت حاوی یک‌لایه تعبیه کلمات[1] از نوع BERT، دو لایه رمزگذار توجه[2]، یک لایه توجه ویژه[3] و یک لایه خروجی[4] تشکیل شده‌است که در ادامه به شرح هر لایه می‌پردازیم:

۳-۱- لایه تعبیهٔ کلمات

مدل‌هایی که باعث می‌شوند که کلمات با کیفیت بالا و قابل خواندن برای سیستم نمایش داده شوند، تعبیه کلمات نام دارند که درواقع کلمات را به بردارهایی از اعداد تبدیل می‌کنند. روش مورد استفاده در این مقاله، روش BERT است.

در این روش، زمینه و هدف به‌صورت زیر نمایش داده می شوند:

"[CLS] + زمینه + [SEP]"

"[CLS] + هدف + [SEP]"

۳-۲- لایهٔ رمزگذار توجه نخست

همان‌طورکه درقبل اشاره شد، از لایه رمزگذار توجه استفاده کردیم. این روش به‌صورت موازی می‌تواند به تحلیل هم‌زمان کلمات ورودی بپردازد؛ بنابراین جایگزین خوبی برای افزایش سرعت تحلیل متن (نسبت به معماری LSTM) به‌شمار می‌رود. در این لایه ابتدا ماژول توجه چندسر(MHA) قرار دارد که نوعی توجه است و توانایی این را دارد که چندین توجه را به‌طور موازی و هم‌زمان انجام دهد. تفاوتی که این ماژول با ترنسفورمرها دارد، این است که اهداف براساس یک زمینهٔ خاص مدل‌سازی می‌شوند. به‌خاطر این تفاوت در هر لایه توجه از یک توجه چندسر برون‌گرا (Inter-MHA) و یک توجه چندسر درون‌گرا (Intra-MHA) برای مدل‌سازی متن و کلمات هدف استفاده شده‌است. برای درک بهتر قبل از تشریح MHA ها به تشریح توجه می پردازیم:

یک توجه[5] دنباله‌ای از کلید(k)[6] و پرس و جو(q)[7]را در یک تابع هم ترازی(f_s) به منظور ایجاد ارتباط معنایی بین q_j و k_i به دنباله خروجی[8] به‌صورت زیر تبدیل می‌کند[۱۳]:

$$k=\{k_1, k_2, \dots, k_n\} \quad (۱)$$

$$q=\{q_1, q_2, \dots, q_m\} \quad (۲)$$

$$\text{Attention}(k, q = \text{softmax}(f_s(k, q))k \quad (۳)$$

$$f_s(k, q) = \tanh([k_i; q_j].W_{att}) \quad (۴)$$

لازم به ذکر است که وزن‌های یادگیری(W_{att}) در ابتدا به‌صورت تصادفی مقداردهی می‌شوند.

۳-۲-۱- توجه چندسر(MHA)

در مقایسه با سازوکار توجه معمولی، MHA می‌تواند محاسبات موازی را روی اطلاعات ورودی انجام دهد و امتیازهای مختلف (n_head) را در زیرفضاهای موازی یاد بگیرد که برای ترازکردن مؤثر است. خروجی های (n_{head}) به هم متصل و به بعد پنهان(d_{head}) نمایش داده می‌شوند که به‌صورت زیر تعریف می‌شود. نکته قابل توجه این است که علامت ; به معنای الحاق بردارها است و $W_{mh} \in R^{d_{hid}*d_{hid}}$ $o^h = \{o_1{}^h, o_2{}^h, \dots, o_m{}^h\}$ خروجی h اُمین توجه سر است[۱۳].

$$\text{MHA}(k, q) = [o^1; o^2; \dots o^{n_{head}}].W_{mh} \quad (۵)$$

$$o^h = \text{Attention}^h(k, q) \quad (۶)$$

در مدل پیشنهاد شده در هر لایه ی رمزگذار توجه از دو ماژول توجه چندسر استفاده شده‌است :

- **توجه چندسر درون‌گرا** (Intra-MHA):

شرایط منحصربه‌فردی است که در آن طبق شکل(۱) سازوکار توجه در این ماژول به این صورت است که q و k با هم برابر است. ما می‌توانیم نمایش متن درون‌نگر c^{intra}را از تعبیه کلمات متن(context) به‌صورت زیر به‌دست آوریم:

$$c^{intra} = \text{MHA}(e^c, e^c) \quad (۷)$$

$$c^{intra} = \{c_1^{intra}, c_2^{intra}, \dots, c_n^{intra}\} \quad (۸)$$

- **توجه چندسر برون‌گرا** (Inter-MHA):

در ماژول دیگر از توجه چندسر برون‌گر(Inter-MHA) استفاده شده که در آن q با k متفاوت است، در این ماژول با توجه به تعبیه کلمات متن و تعبیه جنبه، می‌توانیم بازنمایی هدف t^{inter}, را از طریق فرمول زیر به‌دست آورد[13]:

۳-۲-۲-تبدیل پیچیدگی نقطه‌ای

داده های متنی را بعد از گذر از ماژول MHA می توان با

$$= \text{MHA}(e^c, e^t) \quad t^{inter} \quad (۹)$$

$$t^{inter} = \{t_1^{inter}, t_2^{inter}, \dots, t_m^{inter}\} \quad (۱۰)$$

[1] Word Embeddings
[2] Attentional Encoder Layer
[3] Target-specific Attention Layer
[4] Output Layer
[5] Attention
[6] key
[7] query
[8] output

[DOI: 10.61186/jsdp.20.4.121]

استفاده از یک تبدیل پیچشی نقطه‌ای (PCT) تبدیل کرد. نقطه‌ای نشان می‌دهد که اندازه‌های هسته یک است و تغییر یکسانی برای هر نشانه در ورودی اعمال می‌شود[۱۳]. با توجه به اینکه ما در مدل پیشنهادی از دو ماژول توجه چندسر (MHA) استفاده کرده‌ایم؛ بنابراین از دو ماژول (PCT) به‌ازای خروجی حاصل از هر ماژول قبل استفاده می‌شود؛ درکل می‌توان گفت PCTها برای به‌دست‌آوردن حالت‌های پنهان خروجی لایه رمزگذار توجه از خروجی‌های ماژول چندسر (MHA) یعنی c^{intra}، t^{inter} استفاده می کنند:

$$h^c = PCT(c^{intra}) \quad (۱۱)$$

$$h^c=\{h_1^c, h_2^c, \dots, h_n^c\} \quad (۱۲)$$

$$h^t = PCT(t^{inter}) \quad (۱۳)$$

$$h^t=\{h_1^t, h_2^t, \dots, h_n^t\} \quad (۱۴)$$

تبدیل پیچشی نقطه‌ای(PCT) با توجه به دنباله ورودی(h) به‌صورت زیر تعریف می‌شود[۱۳]:

$$=\sigma(h * w_{pc}^1+b_{pc}^1)* w_{pc}^2+b_{pc}^2 PCT(h) \quad (۱۵)$$

که در آن σ فعال‌سازی ELU است، $*$ عمل‌گر کانولوشن است، $w_{pc}^2 \in R^{d_{hid} \times d_{hid}}$ و $w_{pc}^1 \in R^{d_{hid} \times d_{hid}}$ وزن‌های قابل یادگیری دو هسته کانولوشن هستند، $b_{pc}^2 \in R^{d_{hid}}$ و $b_{pc}^1 \in R^{d_{hid}}$ بایاس‌های دو هسته کانولوشن هستند[۱۳].

۳-۳- لایه رمزگذار توجه دوم

در این لایه خروجی ماژول تبدیل پیچشی نقطه‌ای لایه قبلی که درواقع خروجی لایه رمزگذار توجه نخست است، به‌عنوان ورودی، به این لایه فرستاده می‌شود. روند کار در این لایه مشابه لایه نخست است با این تفاوت که ورودی این لایه، بردار تعبیه کلمات نیست و همان‌طور که گفته شد، خروجی لایه قبل که h^c و h^t است، است.

$$cc^{intra} = MHA(h^c, h^c) \quad (۱۶)$$

$$tt^{inter} = MHA(h^c, h^t) \quad (۱۷)$$

$$h^c = PCT(cc^{intra}) \quad (۱۸)$$

$$h^t = PCT(tt^{inter}) \quad (۱۹)$$

۴-۳- لایه توجه ویژه

پس از به‌دست‌آوردن بازنمایی‌های متن و کلمات که حاصل PCT های لایهٔ توجه دوم بود از لایه توجه ویژه که دارای مجموعه‌ای از پارامترهای خاص خود است، برای به‌دست‌آوردن بازنمایی خاص استفاده می‌کنیم:

$$H^{tsc1} = MHA(h^c, h^t) \quad (۲۰)$$

$$h^{tsc1}=\{h_1^{tsc1}, h_2^{tsc1}, \dots, h_m^{tsc1}\} \quad (۲۱)$$

۵-۳- لایهٔ خروجی

در لایه آخر ما میانگین نمایش‌های نهایی خروجی‌های قبلی را باهم ادغام و به‌عنوان نمایش جامع نهایی$\tilde{O}$ به هم متصل و از یک لایه به‌طورکامل متصل برای نمایش بردار پیوستی در فضای رده‌های C هدف استفاده می‌کنیم. که در آن Y توزیع قطبیت احساسات پیش‌بینی شده‌است، wو b پارامترهای قابل یادگیری هستند[۱۳]:

$$\tilde{O} = [h_{avg}^c, h_{avg}^t, h_{avg}^{tsc1}] \quad (۲۲)$$

$$X= \tilde{w}_o^T \tilde{O} + \tilde{b}_o \quad (۲۳)$$

$$Y= softmax(x) \quad (۲۴)$$

$$= \frac{Exp(x)}{\Sigma_{k=1}^{C} exp\ (x)} \quad (۲۵)$$

۶-۳- آموزش

تابع ضرر، ازدست‌دادن متقاطع[1] تنظیم $\mathcal{L}_{lsr}$ و L2 است که به‌صورت زیر تعریف می‌شود: گفتنی است که در فرمول زیر $\hat{y}$ بردار توزیع احساسات واقعی است که به‌صورت (one_hot) نشان داده می‌شود، y بردار توزیع احساسات پیش‌بینی‌شدهٔ لایه خروجی است، L2 ضریب عبارت منظم‌سازی و Θ مجموعه پارامتر است[13]:

$$\mathcal{L}(\theta) = -\sum_{i=1}^{C} \hat{y}^C \log(y^c) + \mathcal{L}_{lsr} + \lambda \sum_{\theta \in \Theta} \theta^2 \quad (۶)$$

۴- پیاده‌سازی و نتایج

در پیاده‌سازی این مدل از جاساز از قبل آموزش دیده‌شده BERT استفاده شده‌است که ابعاد جاسازی آن برابر سیصد است. برای مقداردهی اولیه وزن‌ها از Glorot استفاده شده‌است. تعداد حالت‌های پنهان روی سیصد ضریب آیتم منظم‌سازی L2 روی 10^{-5} و نرخ حذف تصادفی[2] نیز روی ۰.۱ تنظیم شده‌است. برای به‌روزرسانی تمام پارامترها، از بهینه‌ساز[3] Adam استفاده شده‌است. در پیاده‌سازی این پژوهش از زبان پایتون و کتابخانه‌های پایتورچ و ترنسفورمر در محیط گوگل کولب استفاده شده‌است.

۱-۴- مجموعه‌داده

ما در این مقاله از سه مجموعه‌داده استفاده کرده‌ایم. مجموعه‌دادهای نظرات کاربران در حوزهٔ رستوران‌ها و

[1] Cross-entropy loss
[2] dropout
[3] optimizer

[DOI: 10.61186/jsdp.20.4.121]

لپتاپ‌ها، دو مجموعه‌داده‌ای هستند که از SemEval [14] 2014 Task 4 ، مورد استفاده قرار گرفته‌اند. مجموعه‌داده سوم، مجموعه‌داده توییتر ACL 14 Twitter [۱۵] است که در هر سه مجموعه‌داده، قطبیت احساسات مثبت، خنثی و منفی است که در آن‌ها به‌ازای هر قطبیت به‌ترتیب ۱و۰و۱- اختصاص داده شده‌است. تعداد نمونه‌های آموزش و آزمون و اعتبارسنجی مربوط به هر یک از مجموعه‌داده‌ها در جدول (۱) نشان داده شده‌است.

۴-۲- معیارهای ارزیابی

بعد از فرایند آموزش، مهم‌ترین مسئله برای ارزیابی مدل پیشنهادی ارزیابی مدل آموزش‌دیده است. یکی از معیارهای مورد استفاده در این مقاله معیار دقت[1] است که متداول ترین معیار سنجش کارایی است.

$$\text{Accuracy}=\frac{\text{Tp+TN}}{\text{TP+FP+FN+TN}} \quad (۲۷)$$

TP مثبت صحیح، TN منفی صحیح، FP مثبت کاذب، FN منفی کاذب است. دیگر معیار مورد استفاده F1_Score است که درواقع میانگین وزن‌دار از یادآوری[2] و دقت[3] است. یادآوری(recall) تعداد مثبت‌های بازگردانده‌شده و دقت (precision) تعداد اسناد صحیح بازگردانده‌شده را تعریف می‌کند. میزان به‌دست‌آمده از F1_Score عددی بین صفر و یک است که یک بهترین و صفر بدترین مقدار در این معیار است.

$$\text{Recall}=\frac{\text{Tp}}{\text{TP+FN}} \quad (۲۸)$$

$$\text{Precision}=\frac{\text{Tp}}{\text{TP+FP}} \quad (۲۹)$$

$$\text{F1\_Score}=\frac{2\times(\text{Precision}*\text{Recall})}{(\text{Precision}+\text{Recall})} \quad (۳۰)$$

۴-۳- توصیف آزمایش ها

به نظور بررسی دقیق عملکرد مدل پیشنهادی، آن را با چندین مدل مورد مقایسه و بررسی قرار داده‌ایم .لازم به ذکر است که در اینجا ما فقط از نتایج مدل‌ها برای ارزیابی عملکرد مدل پیشنهادی استفاده کرده‌ایم. به‌جز مدل AEN_BERT سایر مدل‌ها را آزمایش و اجرا نکرده‌ایم در زیر هریک از این مدل‌ها به‌اختصار توضیح داده شده‌است:

- ATAE-LSTM[4][10]: این مدل از سازوکار توجه همراه با شبکه حافظه کوتاه‌مدت بلندمدت استفاده می‌کند[۱۶]. در ATAE_ LSTM که یک مدل طبقه‌بندی است، نمایش تعبیه‌شده کلمات جنبه و نمایش تعبیه‌شده عبارت به‌عنوان ورودی به مدل طبقه‌بندی LSTM متصل می‌شوند و سپس از سازوکار توجه برای تنظیم وزن‌ها استفاده می‌کند. [۱۷].
- RAM[5][18]: از یک LSTM دو طرفه برای به‌دست‌آوردن بردارهای حالت پنهان متنی استفاده می‌کند و از یک شبکه واحد بازگشتی دروازه‌ای برای ترکیب خروجی‌های توجه چندگانه برای نمایش جمله استفاده می‌کند.[۱۹]
- MGAN[6][20]: یک مدل توجه چندمنظوره ارائه می‌کند تا اطلاعات تعاملی (کلی و جزیی) را از بین کلمات جنبه و زمینه به‌دست آورد.[۱۷]
- AEN_BERT[۱۳]: از فناوری هموارسازی برچسب‌ها برای رفع مشکل برچسب‌های غیرقابل اعتماد در RNN استفاده می‌کند و یک شبکه رمزگذار توجه برای رسیدگی به این موضوع کهRNN ها نمی‌توانند به‌صورت موازی پردازش شوند، پیشنهاد می‌کند.[۱۹] گفتنی است، نتایج این مدل حاصل آزمایش و اجرا ماست.

گفتنی است که تعبیه کلمات مدل‌های LSTM، ATAE_ LSTM، RAM، MGAN از نوع Embedding

(جدول-۱): تعداد داده‌های آموزش، آزمون و اعتبارسنجی بر اساس مجموعه‌داده‌ها

(Table-1): Number of training, test and validatio n databased on dataset

	رستوران			لپتاپ			توئیتر		
	مثبت	خنثی	منفی	مثبت	خنثی	منفی	مثبت	خنثی	منفی
آموزش	1740	504	643	800	700	363	1263	2482	1254
آزمایش	728	196	196	341	128	169	173	346	173
اعتبارسنجی	424	164	133	194	170	101	298	645	306

[1] accuracy
[2] recall
[3] precision
[4] Attention-based LSTM
[5] Recurrent Attention on Memory
[6] Multiresolution Graph Attention Network

[DOI: 10.61186/jsdp.20.4.121]

و تعبیه کلمات مدل پیشنهادی و AEN_ BERT از نوع BERT است.

۴-۴- نتایج آزمایش ها

همان‌طورکه درقبل بیان شد، مدل پیشنهادی را با معیارهای دقت و F1_Score بر روی سه مجموعه‌داده ارزیابی کرده‌ایم. همچنین به‌منظور ارائه بهتر نحوه عملکرد مدل، کارکرد این مدل را با چهار مدل از مقاله pang و *دوستان*[۱۷] و یک مدل از مقاله Song و *همکاران*[۱۳] مقایسه می‌کنیم. جدول (۲) معیار دقت به‌دست‌آمده برای هر یک از روش‌ها را نشان می‌دهد. همان‌گونه که در جدول مشاهده می‌شود، در مجموعه‌دادگان لپتاپ مدل پیشنهادی نسبت به بهترین روش، ۴.۲۴ درصد بهتر عمل کرده و در مجموعه‌دادهٔ رستوران با ۰.۳۴ درصد بهترین عملکرد را در بین سایر مدل‌ها دارد.

(جدول-۲): مقایسه روش پیشنهادی با سایر مدل‌ها بر روی مجموعه‌دادگان لپتاپ، رستوران و توئیتر براساس معیار دقت(Accuracy)

(Table-2): Comparison of the proposed method with other models in laptop, restaurant and Twitter datasets based on accuracy criteria

مجموعه‌داده / روش	لپتاپ	رستوران	توئیتر
LSTM	۰.۶۱۴۴	۰.۷۳۰۴	۰.۶۴۷۴
ATAE_ LSTM	۰.۶۰۱۹	۰.۷۳۷۵	۰.۶۸۶۴
RAM	۵۹۵۶.۰	۷۱۵۲.۰	۶۳۵۸.۰
MGAN	۰.۵۸۷۸	۰.۷۱۷۹	۰.۶۳۷۳
AEN_ BERT	۰.۷۵۸۶	۰.۸۰۲۷	**۰.۷۱۲۴**
مدل پیشنهادی	**۰.۷۹۱۵**	**۰.۸۰۵۴**	۰.۶۸۷۹

جدول (۳) معیار F1_Score به‌دست‌آمده را برای هر یک از روش‌ها نشان می‌دهد. همان‌گونه که در جدول مشاهده می‌شود، مدل پیشنهادی ما با اختلاف زیاد نسبت به سایر مدل‌ها در مجموعه‌دادگان رستوران بهترین عملکرد را داشته است. به این معنا که نسبت به بهترین مدل قبل از خود ۵.۹۴ درصد بهتر عمل کرده است.

جدول(۳): مقایسه روش پیشنهادی با سایر مدل‌ها بر روی مجموعه‌دادگان لپتاپ، رستوران و توئیتر براساس معیار F1_Score

(Table-3): Comparison of the proposed method with other models in laptop, restaurant and Twitter datasets based on F1_Score criteria

مجموعه‌دادگان / روش	لپتاپ	رستوران	توئیتر
LSTM	۰.۴۴۰۱	۰.۵۳۳۰	۰.۶۰۵۸
ATAE_ LSTM	۰.۴۹۰۹	۰.۵۷۲۵	۰.۶۵۰۱
RAM	۰.۴۳۰۸	۰.۴۶۵۶	۰.۵۹۹۸
MGAN	۰.۴۲۶۴	۰.۴۹۷۴	۰.۵۸۵۶
AEN_ BERT	۰.۷۲۳۳	۰.۶۷۳۵	**۰.۶۸۳۵**
مدل پیشنهادی	**۰.۷۴۳۰**	**۰.۷۱۴۷**	۰.۶۷۵۳

نتایج به‌دست‌آمده از مقایسه مدل پیشنهادی با مدل‌های دیگر نشان می‌دهد که مدل پیشنهادی براساس معیارF1_Score نسبت به سایر مدل‌های LSTM و RAM، MGAN بسیاربهتر عمل کرده و درکل می‌توان گفت روش پیشنهادی در مقایسه با بیش‌تر مدل‌های موجود در هر دو معیار ارزیابی، به‌نسبه بالاتر و عملکرد بهتری داشته است.

۵- نتیجه گیری

در این مقاله فرضیهٔ ما این بود که با توجه به مدل شبکه‌های تعاملی [24] که حاوی چندین Hop بودند، ما نیز تعداد لایه‌های توجه را افزایش دادیم، در این راستا ما دو و سه و چهار لایه به مدل مرجع اضافه کردیم و با توجه به نتایج حاصل از این لایه‌ها و مقایسه و ارزیابی هریک از آن‌ها به این نتیجه رسیدیم که بهترین عملکرد مدل‌مان با دو لایه حاصل می‌شود؛ بنابراین نوآوری در مدل پیشنهادی افزدون یک لایهٔ دیگر به مدل مرجع[15] بود که دقت مدل پیشنهادی را نسبت به آن افزایش داد.

درکل در این مقاله ما به ارائه یک معماری نوین مبتنی بر یادگیری عمیق برای مسئلهٔ تحلیل احساس مبتنی بر جنبه پرداختیم. برای این هدف ما از شبکه رمزگذار توجه برای طبقه‌بندی احساسات مبتنی بر جنبه استفاده کرده‌ایم که این شبکه برای مدل‌سازی سازوکار توجه بین متن و جنبه مورد استفاده قرار گرفته است. همچنین موضوع غیرقابل اعتمادبودن برچسب را مطرح کرده‌ایم و برای این منظوراز یک رویکرد منظم‌سازی بهره گرفته‌ایم و هموارسازی برچسب (LSR) را به مدل اضافه کرده‌ایم. نتایج ارزیابی مدل پیشنهادی روی سه مجموعه‌داده مختلف، نشان‌دهنده کارآیی بالای معماری پیشنهادی در مقایسه با سایر مدل‌های پیشنهادی در این حیطه است.

۶- مراجع 6-Refrence

[1] Z. Rajabi, M. valavi, and M. Hourali, "Sentiment analysis methods in Persian text: A survey," Signal Data Process., vol. 19, no. 2, pp. 107–132, 2022, doi: 10.52547/jsdp.19.2.107.

[2]"https://daneshyari.com/isi/articles/sentiment_anal."

[DOI: 10.61186/jsdp.20.4.121]

[3] S. Behdenna, F. Barigou, and G. Belalem, “EAI Endorsed Transactions Document Level Sentiment Analysis : A survey,” vol. 4, no. 1, pp. 1–8, 2017.
[4] V. S. Jagtap and K. Pawar, “Analysis of different approaches to Sentence-Level Sentiment Classification,” Int. J. Sci. Eng. Technol., vol. 2, no. 3, pp. 164–170, 2013, [Online]. Available: http://ijset.com/ijset/publication/v2s3/paper11.pdf
[5] H. Wan, Y. Yang, J. Du, Y. Liu, K. Qi, and J. Z. Pan, “Target-aspect-sentiment joint detection for aspect-based sentiment analysis,” AAAI 2020 - 34th AAAI Conf. Artif. Intell., pp. 9122–9129, 2020, doi: 10.1609/aaai.v34i05.6447.
[6] Y. Kim, “Convolutional neural networks for sentence classification,” EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf., pp. 1746–1751, 2014, doi: 10.3115/v1/d14-1181.
[7] A. K. Sharma, S. Chaurasia, and D. K. Srivastava, “Sentimental Short Sentences Classification by Using CNN Deep Learning Model with Fine Tuned Word2Vec,” Procedia Comput. Sci., vol. 167, no. 2019, pp. 1139–1147, 2020, doi: 10.1016/j.procs.2020.03.416.
[8] S. Ramaswamy and N. DeClerck, “Customer perception analysis using deep learning and NLP,” Procedia Comput. Sci., vol. 140, pp. 170–178, 2018, doi: 10.1016/j.procs.2018-.10.326.
[9] H. Sadr, M. mohsen Pedram, and M. Teshnehlab, “Efficient Method Based on Combination of Deep Learning Models for Sentiment Analysis of Text,” Signal Data Process., vol. 19, no. 1, pp. 19–38, 2022, doi: 10.52547/jsdp.19.1.19.
[10] Y. Wang, M. Huang, L. Zhao, and X. Zhu, “Attention-based LSTM for aspect-level sentiment classification,” EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc., pp. 606–615, 2016, doi: 10.18653/v1/d16-1058.
[11] T. Chen, R. Xu, Y. He, and X. Wang, “Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN,” Expert Syst. Appl., vol. 72, pp. 221–230, 2017, doi: 10.1016/j.eswa.2016.10.065.
[12] D. Tang, B. Qin, X. Feng, and T. Liu, “Effective LSTMs for target-dependent sentiment classification,” COLING 2016 - 26th Int. Conf. Comput. Linguist. Proc. COLING 2016 Tech. Pap., pp. 3298–3307, 2016.
[13] Y. Song, J. Wang, T. Jiang, Z. Liu, and Y. Rao, “Targeted Sentiment Classification with Attentional Encoder Network,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 11730 LNCS, pp. 93–103, 2019, doi: 10.1007/978-3-030-30490-4_9.
[14] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, and I. Androutsopoulos, “SemEval-2015 Task 12: Aspect Based Sentiment Analysis,” SemEval 2015 - 9th Int. Work. Semant. Eval. co-located with 2015 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. NAACL-HLT 2015 - Proc., pp. 486–495, 2015, doi: 10.18653/v1/s15-2082.
[15] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, “Adaptive Recursive Neural Network for target-dependent Twitter sentiment classification,” 52nd Annu. Meet. Assoc. Comput. Linguist. ACL 2014 - Proc. Conf., vol. 2, pp. 49–54, 2014, doi: 10.3115/v1/p14-2009.
[16] T. S. Ataei, K. Darvishi, S. Javdan, B. Minaei-Bidgoli, and S. Eetemadi, “Pars-ABSA: an Aspect-based Sentiment Analysis dataset for Persian,” pp. 1–6, 2019.
[17] G. Pang, K. Lu, X. Zhu, J. He, Z. Mo, and Z. Peng, “Aspect-Level Sentiment Analysis Approach via BERT and Aspect Feature Location Model,” vol. 2021, 2021.
[18] Z. Sun, L. Bing, W. Yang, and P. Chen, “Recurrent Attention Network on Memory for Aspect Sentiment Analysis,” pp. 452–461, 2017.
[19] R. Wang, “Interactive Attention Encoder Network with Local Context Features for Aspect-Level Sentiment Analysis,” no. Iccc, pp. 571–576, 2020, doi: 10.1109/ICCC49849.2020.9238924.
[20] F. Fan, Y. Feng, and D. Zhao, “Multi-grained Attention Network for Aspect-Level Sentiment Classification,” pp. 3433–3442, 2018.

سمیه کریمی مدرک کارشناسی خود را در رشته مهندسی رایانه گرایش نرم افزار را از دانشگاه صنعتی قوچان دریافت کرد و در حال حاضر دانشجوی کارشناسی ارشد در گرایش هوش مصنوعی در دانشگاه صنعتی شاهرود است. موضوع پایان‌نامه ایشان "تجزیه و تحلیل احساسات مبتنی بر جنبه با استفاده از یادگیری عمیق" است.

نشانی رایانامه ایشان عبارت است از:

Somayehkarimi6@yahoo.com

فاطمه جعفری‌نژاد در حال حاضر استادیار دانشکده مهندسی کامپیوتر دانشگاه صنعتی شاهرود است. پیش از این ایشان فارغ التحصیل دکترا در گرایش هوش مصنوعی از دانشگاه صنعتی شاهرود هستند و مقطع کارشناسی ارشد را در دانشگاه شهید بهشتی نیز در همان گرایش و همچنین مقطع کارشناسی خود را در دانشگاه تهران در رشته مهندسی رایانه گرایش نرم‌افزار به پایان رسانیده است. زمینه و علائق پژوهشی ایشان یادگیری عمیق، یادگیری ماشین، پردازش متن، روش‌های فرمال است.

نشانی رایانامه ایشان عبارت است از:

jafarinejad@shahroodut.ac.ir

[DOI: 10.61186/jsdp.20.4.121]