

بهبود استخراج ویژگی با استفاده از یک مدل یادگیری عمیق گروهی برای تشخیص موجودیت

الهام پارسایی مهر^{۱*}، مهدی فرتاش^۲ و جواد اکبری ترکستانی^۳

اطلاعات مقاله	چکیده
دریافت مقاله: ۱۴۰۰/۰۴/۱۹ پذیرش مقاله: ۱۴۰۰/۱۱/۳۰	یکی از مراحل اولیه در بیشتر پردازش‌های زبان طبیعی، استخراج موجودیت نامدار از جمله است. در این زمینه تکنیک‌های مختلف مبتنی بر یادگیری ماشین ارائه شده است که بدون نیاز به پیچیدگی‌های استخراج ویژگی دستی، دقت بالاتری از خود نشان داده‌اند. لذا، در این تحقیق ما برای گرفتن ویژگی‌های جمله ورودی از ترکیب دو مدل یادگیری عمیق شامل شبکه عصبی کانولوشن ^۱ و همینطور حافظه کوتاه مدت طولانی ^۲ استفاده می‌کنیم. با استخراج ویژگی‌های محلی کلمات توسط شبکه کانولوشن در کنار ویژگی‌های سراسری، اطلاعات بیشتری از جمله جهت کلاسبندی دقیق‌تر موجودیت‌ها بدست می‌آوریم. ما معماری پیشنهادی مان را روی دو دیتاست CoNLL2003 و ACE05 ارزیابی می‌نماییم و نشان می‌دهیم که افزودن شبکه کانولوشن سطح کلمه باعث استخراج اطلاعات محلی مفیدی از کلمات موجود در جمله می‌شود که منجر به افزایش دقت سیستم می‌گردد. در نهایت، کارایی سیستم را با دیگر رقبا مقایسه می‌نماییم و برتری این معماری نسبت به دیگران گزارش داده می‌شود.
واژگان کلیدی: تشخیص موجودیت نام‌دار، تعبیه‌گذاری کلمه، LSTM، CNN، پردازش زبان طبیعی.	

۱- مقدمه^۳

یکی از گام‌های اولیه در پردازش زبان طبیعی، تشخیص موجودیت نامدار^۴ است. این موضوع توجه بسیاری از محققان را به جهت اهمیت جایگاه این پردازش، به خود جلب کرده است. در این رابطه تکنیک‌های مختلفی ارائه شده است که از روش‌های استخراج ویژگی دستی مانند هجی کلمات و کاپیتالایزن استفاده می‌کنند [۱ و ۲]. از آنجایی که این روش‌ها از دانش خاصی (مثلا دیکشنری اختصاصی) برای اهدافشان استفاده می‌کنند، لذا توسعه آنها به زمینه‌های دیگر کاری پرهزینه و زمانبر است. از طرفی تأثیر کاربرد شبکه‌های عصبی در زمینه‌های مختلف [۳-۷]، محققین را به تحقیقات بیشتر در این عرصه ترغیب نمود. تا بعدها با ارائه شبکه‌های عصبی عمیق موفق به استخراج ویژگی از داده‌های ترتیبی (مانند: صوت، تصویر

و متن) شدند. در سال‌های اخیر، تلاش‌های زیادی برای حل مشکل استخراج موجودیت صورت گرفته و تعدادی راه‌حل ارائه شده که از شبکه‌های عصبی غیرخطی، موسوم به تعبیه گذاری^۵ استفاده می‌کنند. نتایج گزارش شده از این نگرش‌ها روی پروژه‌های پردازش زبان طبیعی کارایی چشمگیرشان را ثابت می‌کند. کُلبِرت و همکارانش یک مدل مبتنی بر شبکه عصبی پیشرو را ارائه دادند که نه تنها به مهندسی ویژگی کمی نیاز دارد، بلکه برای نمایش کلمه از روش تعبیه‌گذاری استفاده می‌کند [۸]. با این وجود، سیستم پیشنهادی‌شان دارای چند محدودیت می‌باشد: (۱) ارتباطات با فاصله میان کلمات در نظر گرفته نمی‌شود. (۲) در مرحله نمایش کلمه، از تعبیه‌گذاری سطح کاراکتر (که

¹ Convolutional Neural Network (CNN)

² long short term memory (LSTM)

⁴ Named Entity Recognition (NER)

⁵ Embedding

* پست الکترونیک نویسنده مسئول: parsayeemehr@gmail.com

۱. عضو هیات علمی، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد

شوشتر

۲. استادیار، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد اراک

۳. استادیار، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد اراک

- گرفتن اطلاعات محلی در کنار اطلاعات سراسری از طریق شبکه کانولوشن سطح کلمه.
- بررسی وجود یا عدم وجود هر مولفه‌ی معماری پیشنهادی روی کارایی سیستم.
- مقایسه دقت معماری مان با مدل‌های مشابه پیشین.

۲- مطالعات پیشین

به طور کلی تلاش‌های انجام شده برای استخراج موجودیت نامدار را می‌توان به سه دسته تقسیم کرد: تکنیک‌های مبتنی بر قاعده [۱۰-۱۳]، تکنیک‌های مبتنی بر روش‌های یادگیری با ناظر [۱۴-۱۷]، و تکنیک‌های مبتنی بر روش‌های یادگیری بدون ناظر [۱۸ و ۱۹].

در همه این نگرش‌ها محاسبات سنگینی برای طراحی ویژگی‌های دستی انجام می‌شود از قبیل: ایجاد دیکشنری، Part Of Speech (POS)، املا کلمات و کاپیتالایزیشن. بنابراین کارایی این سیستمها به شدت تحت تاثیر دقت به کار گرفته شده برای طراحی و مهندسی این ویژگی‌های دستی قرار می‌گیرد.

برای فائق آمدن بر این مسئله‌ی حجم عظیم محاسبات جهت استخراج ویژگی‌های دستی، نگرش شبکه‌های عصبی عمیق ارائه شدند که توسط آنها ویژگی‌های متنی گرفته می‌شوند. قبل از ابداع روش تعبیه گذاری کلمه، پتاسیس و همکاری‌های مدلی ارائه دادند که از شبکه عصبی پیشرو برای استخراج موجودیت استفاده می‌کرد. آنها نتایج چشمگیری را روی دیتاست MUC6 گزارش دادند [۲۰]. علاوه بر این هم‌رئین برای شناسایی موجودیت‌ها از یک نگاهت خود-سازمان برای نمایش توکن‌ها استفاده کرد و LSTM را برای فرموله کردن متن به کار گرفت و نتایج بهتری نسبت به رقبا به دست آورد [۲۱].

بعدها با معرفی تعبیه گذاری کلمه^۳، کلبرت و همکاری‌های یک معماری مبتنی بر شبکه عصبی عمیق پیشنهاد دادند که در آن از تعبیه‌گذاری کلمه و شبکه کانولوشن سطح کلمه برای دو وظیفه تشخیص موجودیت نامدار و POS استفاده می‌شد و نتایج بهتری را گزارش دادند. همچنین هوانگ و همکاری‌های یک LSTM سطح کلمه پیشنهاد کردند [۲۲]؛ آنها ثابت نمودند که افزودن یک لایه CRF

ویژگی‌های سطح کاراکتر را استخراج می‌کند) استفاده نمی‌شود. بنابراین، کلمات ناشناخته (خارج از دیکشنری) که در مرحله آموزش دیده نشده‌اند، به درستی مدیریت نمی‌شوند. ما و هاوی یک سیستم شبکه عصبی ترکیبی ارائه دادند که شبکه عصبی کانولوشن را برای تعبیه سطح کاراکتر به کار می‌گیرد و از ^۱ BLSTM برای دریافت اطلاعات متنی استفاده می‌کند [۹]. بنابراین، ارتباطات بافاصله میان کلمات به خاطر سپرده می‌شود. در پایان، این سیستم از روش CRF^۲ برای دزد کردن مرتبط استفاده می‌کند و به دقت بسیار بالاتری می‌رسد.

در این مقاله ما قصد داریم تا مدل ارائه شده توسط ما و هاوی را از طریق قراردادن یک شبکه کانولوشن سطح کلمه در کنار شبکه BLSTM جهت گرفتن ویژگی‌های بیشتر جمله ورودی توسعه دهیم. ما نشان می‌دهیم که اطلاعات محلی دریافت شده از طریق این شبکه کانولوشن در کنار اطلاعات متنی گرفته شده از طریق BLSTM، منجر به بهبود کارایی سیستم می‌شود. به عبارتی دیگر، در این تحقیق ما یک معماری شبکه عصبی عمیق قدرتمند برای تشخیص موجودیت نامدار ارائه می‌دهیم که به صورت یک سیستم پایانه به پایانه عمل می‌کند که نیازی به مهندسی ویژگی‌های دستی یا دانش خاصی (مانند دیکشنری اختصاصی) ندارد. بنابراین مدل ما قابل استفاده بر روی کاربردهای مشابه دیگر در زمینه‌ها و زبان‌های دیگر هم می‌باشد. در معماری پیشنهادی، ما ابتدا از اتصال خروجی دو مولفه‌ی تعبیه‌گذاری کلمه و تعبیه‌گذاری کاراکتر مبتنی بر شبکه کانولوشن برای نمایش کلمه استفاده می‌کنیم. سپس بردار نمایش به دست آمده را به دو مولفه شبکه کانولوشن سطح کلمه و BLSTM جهت گرفتن اطلاعات محلی و سراسری جمله می‌خورانیم. در نهایت با استفاده از روش CRF، برچسب نهایی را تعیین می‌کنیم. ما کارایی معماری مان را روی دو دیتاست CoNLL2003 و ACE05 بررسی می‌کنیم و برتری آن را به نسبت دیگر رقبا نشان می‌دهیم. اهداف اصلی این مطالعه عبارتند از:

- ارائه یک معماری جدید پایانه به پایانه برای برچسب گذاری توالی کلمات در عمل استخراج موجودیت نامدار.

¹ Bidirectional LSTM

² Conditional Random Field

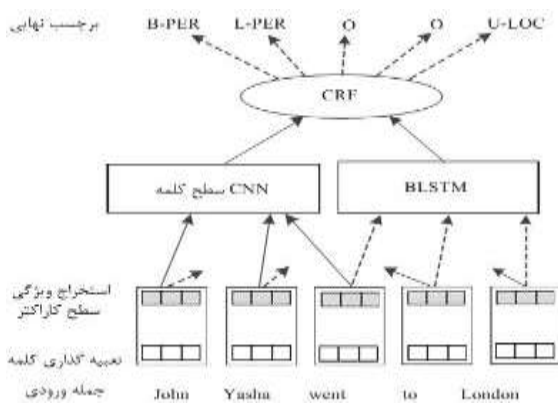
³ word embedding

CHEMDNER در زمینه پزشکی پیاده شد و به معیار اف ۸۳,۷۱ رسیدند [۳۰].

اخیرا تلاشهایی توسط میوا و بنسال و همچنین پتاسیس و همکارانش صورت گرفته که به استخراج توام موجودیت‌ها و رابطه بین آنها از طریق شبکه‌های عصبی عمیق می‌پردازد [۳۱ و ۳۰]. نتایج گزارش شده تأثیر مثبت این نگرش توام را نشان می‌دهد.

۳- معماری پیشنهادی

در این بخش، معماری پیشنهادی‌مان را که بر مبنای ترکیبی از مدل‌های مختلف شبکه‌های عصبی عمیق می‌باشد را معرفی می‌کنیم. همانطور که در شکل (۱) نشان داده شده است، این معماری از پنج مولفه تشکیل شده است: تعبیه‌گذاری کلمه، استخراج ویژگی سطح کاراکتر، شبکه کانولوشن سطح کلمه و CRF. در زیربخش‌های بعدی این مؤلفه‌ها به طور مشروح توضیح داده خواهند شد.



شکل ۱- معماری پیشنهادی

۳-۱- تعبیه‌گذاری کلمه

تعبیه‌گذاری کلمه موثرترین روش برای نمایش کلمه در پردازش‌های زبان طبیعی می‌باشد که اولین بار توسط گلبرت و همکارانش معرفی شد [۸]. از آنجایی که این نوع نمایش ارتباطات معنایی کلمات را در نظر می‌گیرد، کارایی بالاتری نسبت به روش one-hot از خود نشان می‌دهد. در این تحقیق ما از بردار کلمه پیش-آموزش داده شده‌ی دیکشنری Glove، که در دسترس عموم است، برای تشکیل نمایش تعبیه‌گذاری استفاده می‌کنیم. همانطور که در رابطه (۱) نشان داده شده است، با استفاده از یک ضرب ماتریسی، فرم one-hot کلمه ورودی v_i به فرم تعبیه گذاری e_i تبدیل می‌شود.

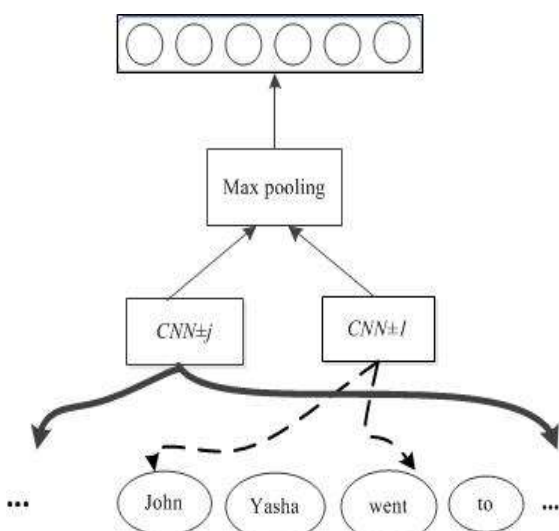
در بالای شبکه عمیق منجر به بهبود کارایی روی دیتاست CoNLL2003 می‌شود. در رابطه با دامنه‌های دیگر، چالپاتی و همکارانش و همچنین زو و همکارانش این معماری شبکه عمیق سطح کلمه را روی اسناد پزشکی به کار گرفتند و دقت قابل توجهی را روی دیتاست تشخیص موجودیت‌های داروها و بیماریها کسب کردند [۲۳ و ۲۴]. علاوه بر این، یان و همکارانش با استفاده از مدل BLSTM عمل تشخیص موجودیت را روی زبان‌های انگلیسی، عربی و آلمانی پیاده نمودند [۲۵].

در تحقیقات اخیر، در کنار شبکه عمیق سطح کلمه، ترکیب شبکه عصبی عمیق سطح کاراکتر با تعبیه کلمه در مرحله نمایش کلمه به کار گرفته شد و دقت بالاتری گزارش داده شد. چپو و نیگل یک شبکه کانولوشن سطح کاراکتر به مدل گلبرت برای نمایش کلمه افزودند و معیار اف ۹۱,۶۲٪ و ۸۶,۲۸٪ روی دیتاست Ontonote5.0 CoNLL2003 را گزارش دادند [۲۳]. همچنین، میساوا و همکارانش این نگرش را برای استخراج موجودیت در زبان ژاپنی به کار گرفتند و نشان دادند که کارایی به دست آمده از دیگر معماری‌های شبکه عمیق به لحاظ دقت پیشی گرفته است [۲۶]. ما و هاوی هم مانند چپو و نیگل یک سیستم پایانه به پایانه پیشنهاد کردند که از تعبیه‌گذاری سطح کاراکتر در مرحله نمایش کلمه کمک می‌گرفت [۹]. هرچند آنها از اطلاعات خارجی استفاده نمی‌کردند. این معماری روی دیتاست CoNLL2003 پیاده شد و معیار اف ۹۱,۲٪ گزارش شد. سنتس و گویماراس با استفاده از نگرشی مشابه، بردار کلمه به دست آمده را به یک شبکه عصبی پیشرو خوراندند و برچسب هر کلمه را از طریق الگوریتم ویتربی پیش‌بینی کردند [۲۸]. آنها معیار اف ۷۱,۲۳ و ۸۲,۲۱ را روی زبان‌های پرتغالی و اسپانیایی از دیتاست CoNLL2002 گزارش دادند. علاوه بر این، لمپل و همکارانش هم تعبیه‌گذاری سطح کاراکتر را با تعبیه گذاری کلمه ادغام کردند؛ ولی آنها از شبکه BLSTM بجای شبکه کانولوشن برای تعبیه کاراکتر استفاده نمودند [۲۹]. آنها مدل‌شان را برای استخراج موجودیت روی دیتاست CoNLL2002 پیاده نمودند و به ترتیب مقادیر معیار اف ۷۸,۷۶، ۹۰,۹۴، ۸۱,۷۴ و ۸۵,۷۵ را روی زبان‌های آلمانی، انگلیسی هلندی و اسپانیایی به دست آوردند. همچنین این مدل توسط حبیبی و همکارانش روی دیتاست

به طوری که $f^k \in R^{n-w+1}$ یک نگاشت ویژگی است، w عرض فیلتر است و n معرف طول کلمه ورودی k می‌باشد. در پایان، بردار جمله ورودی در نمایش سطح کاراکتر، $chr_s = \{y_1, y_2, \dots, y_l\} \in R^{l \times d}$ ، به مؤلفه بعدی داده می‌شود. بطوری که d ابعاد تعبیه‌گذاری کاراکتر می‌باشد و l تعداد کاراکترها را در جمله ورودی تعیین می‌کند.

۳-۳- شبکه کانولوشن سطح کلمه

با توجه به شکل (۱)، بردار خروجی دو مؤلفه تعبیه‌گذاری کلمه و نمایش سطح کاراکتر با یکدیگر ادغام شده و بردار ترکیبی $[emb_s, chr_s]$ به دست آمده به این مؤلفه به عنوان ورودی داده می‌شود. در این مؤلفه با استفاده از تکنیک Mix-CNN [۳۲]، اطلاعات محلی کلمات جمله ورودی استخراج می‌شوند. با توجه به شکل (۳)، با استفاده از کلمات همسایه (مابعد و ماقبل) این اطلاعات محلی به دست می‌آیند.



شکل ۳- ساختار mixture CNN (Mix-CNN)

همانطور که در این شکل دیده می‌شود، در $CNN+1$ فقط از یک کلمه قبل تا یک کلمه بعد از کلمه مورد نظر برای استخراج ویژگی‌های معنایی ملاحظه می‌شود (مثلاً برای کلمه Yasha، از کلمه John تا کلمه went). به طور کلی، از J کلمه قبل (سمت چپ) تا J کلمه بعد (سمت راست) کلمه مورد نظر را بررسی می‌کند. ویژگی شناسایی شده توسط i امین فیلتر از شبکه $CNN+J$ را Ze_j^i می‌نامیم. بنابراین مجموعه ویژگی‌های استخراج شده توسط ne فیلتر $E1_j =$ Mix-CNN برای موجودیت $e1$ را به صورت $E1_j = [Ze_j^1, Ze_j^2, \dots, Ze_j^{ne}]$ تعریف می‌کنیم. با توجه به این

¹ Max-pooling

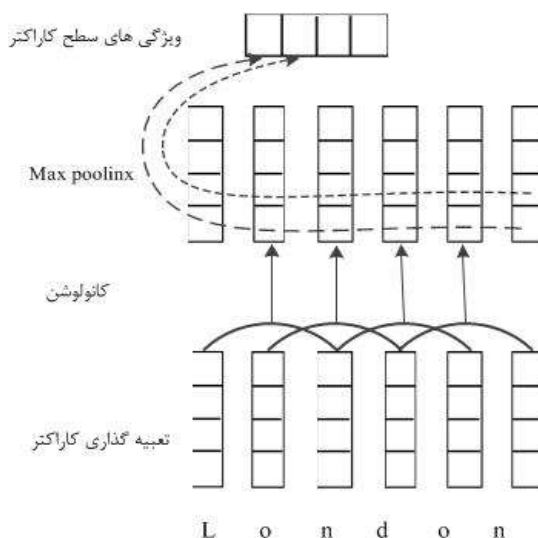
$$e_i = w^{word} v_i \tag{1}$$

w^{word} معرف دیکشنری Glove می‌باشد بنابراین با داشتن l کلمه جمله ورودی، بردار $emb_s = \{e_1, e_2, \dots, e_l\} \in R^{l \times d}$ را تشکیل می‌دهیم؛ بطوریکه d ابعاد تعبیه‌گذاری را نشان می‌دهد. بردار emb_s برای پردازش‌های بیشتر به مؤلفه‌های بعدی داده می‌شود.

۳-۲- استخراج ویژگی سطح کاراکتر

شبکه عصبی عمیق کانولوشن نه تنها در عرصه پردازش تصویر به نتایج چشمگیر دست یافته، بلکه در زمینه پردازش زبان طبیعی هم کارایی موثری از خود نشان داده است. در این مؤلفه از معماری پیشنهادی، یک شبکه عمیق CNN برای استخراج اطلاعات از توالی کاراکترهای ورودی به کار می‌گیریم. از آنجایی که در برخورد با کلمات ناشناخته، نگرش سطح کاراکتر موثرتر از نگرش سطح کلمه عمل می‌کند، ما این مؤلفه را جهت افزایش دقت معماری مان به سیستم می‌افزاییم. همانطور که در شکل (۲) نشان داده شده، در این مؤلفه ابتدا هر کاراکتر از جمله ورودی به یک بردار تعبیه‌گذاری شده تبدیل می‌شود. سپس، این توالی بردارها توسط فیلترهایی (هسته‌ها) با اندازه‌های متفاوت (۳، ۵، ۷) برای دریافت اطلاعات محلی موثر اسکن می‌شوند. بعد یک عمل ماکزیمم‌گیری^۱، که در رابطه (۲) نشان داده شده اعمال می‌شود تا ویژگی‌های با ارزش‌تر انتخاب شوند و پیچیدگی محاسبات بعدی کاهش یابد.

$$y_k = \max(f^k(i)) \tag{2}$$

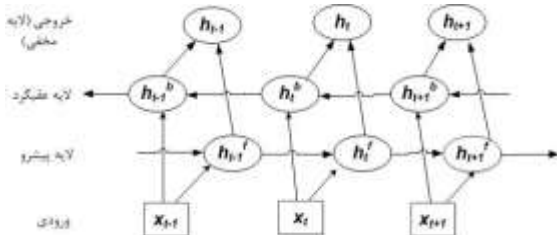


شکل ۲- استخراج ویژگی سطح کاراکتر [۹]

$$h_t^f \text{ or } h_t^b = o_t \odot \tanh(c_t) \quad (10)$$

بطوریکه σ معرف تابع لاجستیک و \odot به عنوان ضرب داخلی می‌باشد. w و u ماتریس‌های وزن و b بردار بایاس است.

با توجه به شکل (۳)، بردار مخفی در لایه پیشرو (h_t^f) اطلاعات را از گذشته (x_1) تا کلمه جاری (x_t) فرموله می‌کند. در مقابل، بردار مخفی در لایه عقبگرد (h_t^b) اطلاعات را از آینده (x_t) تا کلمه جاری (x_t) ارائه می‌دهد. این دو بردار مخفی به یکدیگر متصل می‌شوند تا $h_t = [h_t^f; h_t^b]$ را تشکیل دهد و به ماجول بعدی داده شود.



شکل ۴- ساختار شبکه LSTM دو طرفه

۳-۵-CRF

در عمل استخراج موجودیت نامدار، تخمین برچسب هر کلمه به مقدار برچسب کلمات همسایه‌اش بسیار وابسته است. برای مثال، در طرح BILOU، که در آن برچسب هر موجودیت معرف نوع کلاس آن موجودیت و همچنین موقعیت مکانی آن کلمه در محدوده (عبارت) موجودیت شناسایی شده (Begin, Inside, Last, Outside, Unit) می‌باشد، برچسب I-PER نمی‌تواند بعد از برچسب I-LOC بیاید، ولی می‌تواند بعد از برچسب‌های I-PER یا B-PER بیاید. بنابراین با در نظر گرفتن ارتباط مابین برچسب‌های خروجی، بهتر است برچسب توالی کلمات به صورت مشارکتی و توأم صورت گیرد بجای اینکه به طور مستقل از هم انجام شود. لذا، ما از روش CRF که اولین بار توسط لَفرتی و همکارانش [۳۴] معرفی شد برای تخمین توالی کلمات استفاده می‌کنیم.

در این بخش، ما بردار $q = [E_s, h_t]$ که از الحاق دو بردار ویژگی حاصل از مولفه‌های شبکه کانولوشن سطح کلمه و BLSTM بدست آمده را تشکیل می‌دهیم؛ بطوریکه $q_i \in \mathcal{Y}$ نمایانگر i امین کلمه از جمله ورودی $\mathcal{Y} = \{y_1, y_2, \dots, y_L\}$ می‌باشد. CRF یک توالی از برچسب های خروجی برای کلمه q می‌باشد.

واقعیت که کلمات مختلف وابستگی‌های متفاوتی به کلمات اطرافشان دارند؛ لذا با توجه به رابطه (۳) ما یک عمل ماکزیمم‌گیری روی این ویژگی‌های کسب شده برای انتخاب بهترین‌ها اعمال می‌کنیم تا کارایی بهبود یابد و همچنین محاسبات بعدی نیز کاهش پیدا کنند.

$$EI_s = \begin{pmatrix} \max(Ze_1^1 Ze_2^1 \dots Ze_j^1) \\ \dots \\ \max(Ze_1^n Ze_2^n \dots Ze_j^n) \end{pmatrix} \quad (3)$$

ما بردار به دست آمده را به عنوان ویژگی‌های استخراج شده برای l کلمه جمله ورودی به لایه بعدی می‌دهیم. مولفه‌های بعدی داده می‌شود.

۳-۴-LSTM دو طرفه (BLSTM)

شبکه LSTM دو طرفه اولین بار توسط گریو و سمیدیهیر [۳۱] ارائه شد که در آن توالی کلمات ورودی در هر دو جهت پیشرو و عقبگرد (چپ به راست و راست به چپ) اسکن می‌شوند. بنابراین، نه تنها مانند LSTM یکطرفه اطلاعات از حالات گذشته تا کنون (پیشرو) به دست می‌آیند؛ بلکه اطلاعات از حالات آینده (عقبگرد) نیز به دست می‌آیند. بردار کلماتی که به این شبکه خوراند می‌شود از ترکیب خروجی‌های دو مولفه تعبیه‌گذاری کلمه و نمایش سطح کاراکتر به دست می‌آید. لذا با توجه به رابطه (۴) دو بردار حاصل شده از دو مولفه ذکر شده با هم ادغام شده تا بردار x_t شکل گیرد.

$$x_t = [emb_s \oplus chr_s] \quad (4)$$

به طوریکه علامت \oplus عمل الحاق را نشان می‌دهد. در این شبکه، یک واحد LSTM در گام زمانی t از طریق روابط (۵-۱۰) به روز می‌شوند. بطوریکه سلول حافظه c_t و گیت ورودی i_t ، گیت خروجی o_t ، گیت فراموشی f_t و بردار مخفی با ابعاد n_{lt} از طریق این محاسبات بدست می‌آیند:

$$i_t = \sigma(w^{(i)}x_t + U^{(i)}h_{t-1} + b^{(i)}) \quad (5)$$

$$f_t = \sigma(w^{(f)}x_t + U^{(f)}h_{t-1} + b^{(f)}) \quad (6)$$

$$o_t = \sigma(w^{(o)}x_t + U^{(o)}h_{t-1} + b^{(o)}) \quad (7)$$

$$u_t = \tanh(w^{(u)}x_t + U^{(u)}h_{t-1} + b^{(u)}) \quad (8)$$

$$c_t = i_t \odot u_t + f_t \odot c_{t-1} \quad (9)$$

جدول ۱- لیست مقادیر ابرپارامترهای مورد استفاده

نام ابرپارامتر	توضیحات	مقدار
D	ابعاد تعبیه‌گذاری	۳۰۰
Ne	تعداد فیلترهای کانولوشن سطح کلمه	۱۰۰
n_lt	تعداد واحدهای مخفی BLSTM	۱۰۰
J	تعداد شبکه کانولوشن در Mix-CNN	۳
ρ	dropout	۰,۵
	Epoch	۵۰
	Learning rate	۰,۰۱۵
	Batch size	۱۰

۴-۳- نتایج

ما در ابتدا میزان کارایی وجود مولفه‌های مختلف معماری مان را ارزیابی می‌کنیم. سیستم‌هایی که به جهت این ارزیابی به مقایسه آنها می‌پردازیم، به شرح زیر می‌باشند:

- BLSTM: در این سیستم فقط از شبکه BLSTM برای دریافت اطلاعات جمله استفاده می‌شود.
- BLSTM + CNN: در این سیستم از ترکیب شبکه کانولوشن سطح کلمه با BLSTM برای دریافت اطلاعات محلی و سراسری استفاده می‌شود.
- BLSTM + CNN + CRF: در این سیستم یک لایه CRF به معماری BLSTM + CNN اضافه می‌شود.

همانطور که در جدول ۲ دیده می‌شود، با افزودن یک کانولوشن سطح کاراکتر در کنار تعبیه‌گذاری کلمه برای نمایش کلمه، کارایی سیستم برای دیتاست‌های CoNLL2003 و ACE05 به ترتیب به مقادیر ۱,۸۱ و ۱,۵۱ افزایش یافته است. مدل BLSTM+CNN نسبت به مدل BLSTM دقت بالاتری در هر دو دیتاست نمایش می‌دهد که نشان‌دهنده موثر بودن ایده ما در این تحقیق (ترکیب شبکه کانولوشن سطح کلمه با BLSTM) می‌باشد. در پایان مشاهده می‌شود که با افزودن یک لایه CRF (ردیف آخر جدول)، مقدار معیاراف افزایش می‌یابد.

۴-۴- بحث

همانطور که جدول ۳ نشان می‌دهد ما کارایی معماری پیشنهادی را با دیگر سیستم‌ها مقایسه نمودیم که عبارتند از: مدل BLSTM-CRF ارائه شده توسط هوانگ و همکارانش

CRF با استفاده از رابطه (۱۱) یک احتمال شرطی $p(y|q; W, b)$ برای y با شرط q داده شده، را تعریف می‌کند.

$$p(y|q; W, b) = \frac{\prod_{i=1}^n \psi_i(y_{i-1}, y_i, q)}{\sum_{y' \in \mathcal{Y}(q)} \prod_{i=1}^n \psi_i(y'_{i-1}, y'_i, q)} \quad (11)$$

بطوریکه $\mathcal{Y}(q)$ توالی از برجسبهای ممکن برای q را نشان می‌دهد. $\psi_i(y', y, q) = \exp(W_{y', y}^T z_i + b_{y', y})$. توابع پتانسیل هستند؛ $W_{y', y}^T$ و $b_{y', y}$ به ترتیب بردارهای وزن و بایاس برای جفت برجسب (y', y) می‌باشند. همانطور که در رابطه (۱۲) نشان داده شده، از احتمال منفی به عنوان تابع زیان در مرحله آموزش استفاده شده است.

$$L(W, b) = -\sum_i \log p(y; W, b) \quad (12)$$

۴-آزمایشات

ما کارایی معماری شبکه عمیق پیشنهادی را روی دو دیتاست عمومی ارزیابی می‌کنیم. در زیربخش‌های بعدی جزئیات دیتاست‌ها و همینطور مقادیر ابرپارامترهای مورد استفاده در پیاده‌سازی مان را به طور مشروح معرفی می‌کنیم.

۴-۱- دیتاست

ما آزمایشات را روی دو دیتاست عمومی CoNLL2003 و ACE05 پیاده نمودیم. در دیتاست CoNLL2003، چهار نوع موجودیت وجود دارد که عبارتند از: نام‌های سازمان‌ها، مکان‌ها، اشخاص و غیره. در دیتاست ACE05، هفت نوع موجودیت داریم که شامل: نام‌های سازمان‌ها، مکان‌ها، اشخاص، وسایل نقلیه، موجودیت‌های ژئو-سیاسی، امکانات و اسلحه.

ما از طرح BILOU برای پروسه تشخیص موجودیت استفاده کردیم که در آن برجسب هر موجودیت معرف نوع کلاس آن موجودیت و همچنین موقعیت مکانی آن کلمه در محدوده (عبارت) موجودیت شناسایی شده (Begin, Inside, Last, Outside, Unit) می‌باشد.

۴-۲- ابرپارامترها

ما از زبان برنامه‌نویسی پایتون و توابع کتابخانه‌ای کراس برای پیاده‌سازی معماری پیشنهادی استفاده کردیم. لیست مقادیر ابرپارامترهای مورد استفاده در آزمایشاتمان در جدول ۱ نشان داده شده است.

همکارانش [۳۵]، یک سیستم پایانه به پایانه توسط میوا و بنسال [۳۱]، و یک مدل توأم توسط پارسی‌مهر و همکارانش [۲].

[۲۲]، نگرش LSTM-CNN معرفی شده توسط چپو و نیگل [۲۶]، مدل BLSTM-CRF پیشنهاد شده توسط لمپل و همکارانش [۲۹]، BLSTM-CNNs-CRF توسط ما و هاوی [۹]، یک نگرش مشارکتی توسط ژنگ و

جدول ۲- نتایج کارایی معماری‌های مختلف. "WE" یعنی از تعبیه‌گذاری کلمه برای نمایش کلمه استفاده شده و "char" یعنی از CNN سطح کاراکتر نیز به مدل جهت نمایش کلمه افزوده شده است.

نام مدل	فرم نمایش کلمه	CoNLL2003			ACE05		
		دقت	فراخوانی	معیار اف	دقت	فراخوانی	معیار اف
BLSTM	WE	۸۵٫۵۶	۸۵٫۰۱	۸۵٫۲۸	۸۰٫۸۳	۸۲٫۲۳	۸۱٫۶۲
BLSTM	WE + char	۸۷٫۳۳	۸۶٫۸۵	۸۷٫۰۹	۸۲٫۷۹	۸۳٫۱۳	۸۳٫۱۳
BLSTM+CNN	WE + char	۹۰٫۵	۸۹٫۹۶	۹۰٫۲۳	۸۳٫۵۵	۸۴٫۹۳	۸۴٫۲۳
BLTM+CNN+CRF	WE + char	۹۱٫۷۸	۹۱٫۲۹	۹۱٫۲۳	۸۴٫۹۴	۸۵٫۶۲	۸۵٫۲۸

مشارکت شبکه کانولوشن) برای گرفتن اطلاعات سطح کاراکتر و سطح کلمه استفاده می‌نمود.

سیستم BLSTM-CNN-CRF ارائه شده توسط ما و هاوی، که شبیه‌ترین مدل به معماری ما می‌باشد؛ از شبکه کانولوشن برای نمایش اطلاعات سطح کاراکتر و از BLSTM برای نمایش اطلاعات متنی استفاده نمود و در لایه آخر CRF را برای برچسب‌گذاری توالی کلمات بکار گرفت. معماری ما از لحاظ مشارکت شبکه کانولوشن برای نمایش اطلاعات سطح کلمه با این نگرش متفاوت است.

علاوه بر تحقیقات ذکر شده که بر روی دیتاست CoNLL2003 برای استخراج موجودیت انجام شد، میوا و بنسال و همچنین ژنگ و همکارانش در کارهای جداگانه‌ای به استخراج همزمان موجودیت‌ها و روابط بین آنها روی دیتاست ACE05 پرداختند. دو مدل ارائه شده این پژوهشگران تفاوت‌هایی با معماری ما دارند: (۱) عدم اعمال لایه CRF. (۲) عدم مشارکت شبکه کانولوشن. در تحقیق صورت گرفته توسط پارسی‌مهر و همکارانش [۲]، یک سیستم توأم برای شناسایی موجودیت‌ها و نوع رابطه بین آنها ارائه شده که مولفه استخراج موجودیت در این سیستم از چند نظر با مدل پیشنهادی این تحقیق تفاوت دارد: (۱) عدم استفاده از CNN سطح کاراکتر برای نمایش کلمه. (۲) عدم استفاده از لایه CRF. گرچه در جدول ۳ دیده می‌شود که نتایج بدست آورده شده توسط لیو و همکارانش [۳۶] و همچنین توسط زیا و همکارانش [۳۷] بر روی دیتاست CoNLL03 از نتایج ما بهتر است، ولی به علت اینکه در این دو تحقیق از دانش خارجی برای شناسایی و کلاس بندی موجودیت‌ها استفاده کردند؛ لذا به عنوان یک سیستم

جدول ۳- مقایسه معماری ما با دیگر سیستم‌های مشابه

نام مدل	معیار اف	
	CoNLL-2003	ACE05
هوانگ و همکارانش	۹۰٫۱۰	-
چپو و نیگل	۹۰٫۷۷	-
لمپل و همکارانش	۹۰٫۹۴	-
ما و هاوی	۹۱٫۲۱	-
میوا و بنسال		۸۳٫۴
ژنگ و همکارانش		۸۴٫۲
پارسی‌مهر و همکاران	-	۸۳٫۱
لیو و همکارانش	۹۲٫۷۵	-
زیا و همکارانش	۹۲٫۲۵	۷۸٫۲
معماری ما	۹۱٫۵۳	۸۵٫۲۸

در مدل BLSTM-CRF ارائه شده توسط هوانگ و همکارانش، یک شبکه عصبی BLSTM برای نمایش اطلاعات سطح کلمه بکار برده شده و بر خلاف معماری ما از هیچکدام از شبکه‌های کانولوشن سطح کلمه و سطح کاراکتر برای نمایش اطلاعات استفاده نشد. از آنجایی که به این مدل ویژگی‌های دستی داده می‌شود (مانند هجی کلمات)؛ نمی‌توان آن را به عنوان یک معماری پایانه به پایانه در نظر گرفت. چپو و نیگل مدلی را معرفی کردند که از شبکه کانولوشن سطح کاراکتر برای نمایش اطلاعات کلمه و از BLSTM برای مدل‌سازی ویژگی‌های سطح کلمه استفاده می‌نمود. این معماری دو تفاوت عمده با معماری ما دارد: (۱) از شبکه کانولوشن سطح کلمه برای نمایش اطلاعات متنی استفاده نشد. (۲) لایه CRF بکار گرفته نشد. مدل پیشنهادی لمپل و همکارانش از BLSTM (بدون

عصبی عمیق برای استخراج موجودیت‌های نامدار ارائه دادیم که بدون نیاز به دانش خاصی (مثلا دیکشنری اختصاصی) مانند یک سیستم پایانه به پایانه عمل می‌کند. بنابراین مدل ما قابلیت اجرا روی زبان‌ها و وظایف دیگر را دارد. علاوه بر این ما از ترکیب دو مولفه تعبیه‌گذاری کلمه و شبکه کانولوشن سطح کاراکتر جهت نمایش یک بردار کلمه استفاده نمودیم. همانطور که جدول ۲ نشان داد، کارایی وجود مولفه‌های مختلف معماری‌مان را ارزیابی نمودیم. همچنین، کارایی سیستم‌مان را با مدل‌های مشابه دیگر مقایسه نمودیم. با توجه به جدول ۳ مشاهده می‌شود که معماری پیشنهادی از لحاظ معیار اف از دیگر رقبا برتری دارد.

در آینده قصد داریم تا روی متون زبان فارسی تمرکز کنیم و عمل تشخیص موجودیت را روی این زبان پیاده کنیم. از آنجایی که مدل پیشنهادی ما به زبان خاصی وابسته نیست، لذا رسیدن به این هدف کار سختی نخواهد بود.

پایانه-به-پایانه در نظر گرفته نمی‌شوند. بنابراین میتوان گفت که معماری پیشنهادی از نظر معیار اف روی هر دو دیتاست CoNLL2003 و ACE05 از دیگر رقبا پایانه-به-پایانه پیشی می‌گیرد.

۴- نتیجه گیری

در این تحقیق، ما یک قالب کاری بر مبنای روش‌های یادگیری عمیق برای تشخیص موجودیت‌های نامدار ارائه دادیم که اطلاعات مؤثرتری را از جمله ورودی استخراج می‌کند و همچنین به عملیات مهندسی ویژگی احتیاجی ندارد. به بیان دقیق‌تر، ما مدل ارائه‌شده توسط ما و هاوی را با افزودن یک شبکه کانولوشن سطح کلمه در کنار BLSTM برای گرفتن اطلاعات متنوع‌تری بهبود دادیم. ما نشان دادیم که اطلاعات محلی استخراج شده توسط این شبکه کانولوشن در کنار اطلاعات متنی دریافت شده توسط BLSTM منجر به بهبود کارایی می‌شود. به عبارتی دیگر، در این مطالعه، ما یک معماری قدرتمند بر پایه شبکه‌های

مراجع

[1] A. Akkasi, and E. Varoglu, "Improvement of Chemical Named Entity Recognition through Sentence-based Random Under-sampling and Classifier Combination", Journal of AI and Data Mining, Vol. 7, NO. 2, 2019, pp. 311-319.

[2] E. Parsaeimehr, M. Fartash, and J. Akbari Torkestani, "An Enhanced Deep Neural Network-Based Architecture for Joint Extraction of Entity Mentions and Relations", Journal of Fuzzy logic and Intelligent Systems, Vol. 20, NO. 1, 2020, pp. 69-76.

[۳] علی سلیمانی ایوری، محمد رضا فدوی امیری و حسین مروی، "تولید سیگنال مصنوعی زلزله به کمک مدلی جدید در فشرده سازی و آموزش شبکه های عصبی مصنوعی"، مجله مدل سازی در مهندسی، دانشگاه سمنان، دوره ۱۴، شماره ۴۶، سال ۱۳۹۵، صفحه ۷۵-۸۵.

[۴] علی نظری، "مدلسازی انرژی ضربه ی فولادهای مرتبه ای با استفاده از شبکه های عصبی مصنوعی"، مجله مدل سازی در مهندسی، دانشگاه سمنان، دوره ۱۴، شماره ۴۵، تابستان ۱۳۹۵، صفحه ۱۴۵-۱۶۲.

[۵] زهرا مروج و جواد آذرخش، "شبیه سازی و طبقه بندی وقایع کیفیت توان با استفاده از شبکه عصبی"، مجله مدل سازی در مهندسی، دانشگاه سمنان، دوره ۱۳، شماره ۴۱، تابستان ۱۳۹۴، صفحه ۱۳۷-۱۴۶.

[6] M.A. Spalenza, L. Lusquino-Filho, F. M. G. Franca, P. M. V. Lima, and E. Oliveira1, "LCAD - UFES at FakeDeS 2021: Fake News Detection Using Named Entity Recognition and Part-of-Speech Sequences", *IberLEF@SEPLN 2021*, 2021.

[7] B. Song, F. Li, Y. Liu, and X. Zeng, "Deep learning methods for biomedical named entity recognition: a survey and qualitative comparison", *Briefings in Bioinformatics*, 2021.

[8] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *The Journal of Machine Learning Research*, Vol. 12, 2011, pp. 2493–2537.

[8] X. Ma, and E. Hovy, "End-to-end sequence labeling via bi-directional lstm-cnns-crf", 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 2016.

[10] O. Etzioni, M. Cafarella, D. Downey, A.M. Popescu, T. Shaked, S. Soderland, D.S. Weld, and A. Yates, "Unsupervised named entity extraction from the web: An experimental study", *Artificial intelligence*, Vol. 165, NO. 1, 2005, pp. 91–134.

- [11] S. Zhang, and N. Elhadad, "Unsupervised biomedical named entity recognition: Experiments with clinical and biological texts," *Journal of Biomedical Informatics*, Vol. 46, NO. 6, 2013, pp. 1088–1098.
- [12] A.P. Quimbaya, A.S. Múnera, R.A.G. Rivera, J.C.D. Rodríguez, O.M.M. Velandia, A.A.G. Peña, and C. Labbé, "Named entity recognition over electronic health records through a combined dictionary-based approach", *Procedia Computer Science*, Vl. 100, 2016, pp. 55–61.
- [13] D. Hanisch, K. Fundel, H.T. Mevissen, R. Zimmer, and J. Fluck, "Prominer: rule-based protein and gene entity recognition", *BMC bioinformatics*, Vol. 6, NO. 1:S14, 2005.
- [14] G. Zhou, and J. Su, "Named entity recognition using an hmm based chunk tagger", 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, 2002.
- [15] W. Liao, and S. Veeramachaneni, "A simple semi-supervised algorithm for named entity recognition", *NAACL HLT 2009 Workshop on Semi-supervised Learning for Natural Language Processing*. Boulder, Colorado, 2009.
- [16] J. Hoffart, M.A. Yosef, I. Bordino, H. Fürstenu, M. Pinkal, M. Spaniol, B. Taneva, S. Thater, and G. Weikum, "Robust disambiguation of named entities in text", 2011 Conference on Empirical Methods in Natural Language Processing, Edinburgh, Scotland, UK, 2011.
- [17] Z. Ji, A. Sun, G. Cong, and J. Han, "Joint recognition and linking of fine-grained locations from tweets", *International World Wide Web Conference Committee*, Montréal, Québec, Canada, 2016.
- [18] L. Jafar Tafreshi, and F. Soltanzadeh, "A Novel Approach to Conditional Random Field-based Named Entity Recognition using Persian Specific Features", *Journal of AI and Data Mining*, Vol. 8, NO. 2, 2020, pp. 227-236.
- [19] D. Nadeau, and S. Sekine, "A survey of named entity recognition and classification", *Lingvisticae Investigationes*, Vol. 30, NO. 1, 2007, pp. 3-26.
- [20] G. Petasis, A. Cucchiarelli, P. Velardi, G. Paliouras, V. Karkaletsis, and C.D. Spyropoulos, "Automatic adaptation of proper noun dictionaries through cooperation of machine learning and probabilistic methods", 23rd annual international ACM SIGIR conference on Research and development in information retrieval, Athens, Greece, 2000.
- [21] J. Hammerton, "Named entity recognition with long short-term memory", *Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, 2003, pp. 172–175.
- [22] Z. Huang, W. Xu, and K. Yu, "Bidirectional lstm-crf models for sequence tagging", *arXiv preprint*, 2015. arXiv:1508.01991.
- [23] R. Chalapathy, E. Zare Borzeshi, and M. Piccardi, "An investigation of recurrent neural architectures for drug name recognition", *Seventh International Workshop on Health Text Mining and Information Analysis*, Austin, TX, 2016.
- [24] K. Xu, Z. Zhou, T. Hao, and W. Liu, "A bidirectional lstm and conditional random fields approach to medical named entity recognition", *International Conference on Advanced Intelligent Systems and Informatics*, 2017.
- [25] S. Yan, C. Hardmeier, and J. Nivre, "Multilingual named entity recognition using hybrid neural networks", *Sixth Swedish Language Technology Conference (SLTC)*, 2016.
- [26] J. PC Chiu, and E. Nichols, "Named entity recognition with bidirectional lstm-cnns", *Transactions of the Association for Computational Linguistics*, Vol. 4, 2016, pp. 357-370.
- [27] S. Misawa, M. Taniguchi, Y. Miura, and T. Ohkuma, "Character-based bidirectional lstm-crf with words and characters for Japanese named entity recognition", *First Workshop on Subword and Character Level Models in NLP*, Copenhagen, Denmark, 2017.
- [28] C.N. dos Santos, and V. Guimaraes, "Boosting named entity recognition with neural character embeddings", *Proceedings of the Fifth Named Entity Workshop of the Association for Computational Linguistics*, 2015, pp. 25-33.
- [29] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer. "Neural architectures for named entity recognition", *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2016, pp. 260-270.
- [30] M. Habibi, L. Weber, M. Neves, D.L. Wiegandt, and U. Leser, "Deep learning with word embeddings improves biomedical named entity recognition", *Bioinformatics*, Vol. 33, No. 14, 2017, pp. 137–148.
- [31] M Miwa, and M. Bansal. "End-to-end relation extraction using LSTMs on sequences and tree structures",

Association for Computational Linguistics, 2016, pp. 1105-1116.

[32] S. Zheng, J. Xu, P. Zhou, H. Bao, Q. Zhenyu, and B. Xu, "A neural network framework for relation extraction: learning entity semantic and relation pattern", *Knowledge-Based Systems*, Vol. 114, 2016, pp. 12-23.

[33] A. Graves, and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM networks", *Journal of Neural Networks*, Vol. 18, NO. 5-6, 2005, pp. 602-610.

[34] J. Lafferty, A. McCallum and F. CN Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data", In: *Proceedings of the 18th International Conference on Machine Learning 2001*, San Francisco, US, 2001.

[35] S. Zheng, Y. Hao, D. Lu, H. Bao, J. Xu, H. Hao, and B. Xu. "Joint entity and relation extraction based on a hybrid neural network", *Journal of Neurocomputing*, Vol. 257, 2017, pp. 59-66.

[36] T. Liu, J. Yao, and C. Lin, "Towards improving neural named entity recognition with gazetteers", in *ACL*, 2019, pp. 5301-5307.

[37] C. Xia, C. Zhang, T. Yang, Y. Li, N. Du, X. Wu, W. Fan, F. Ma, and P. S. Yu, "Multi-grained named entity recognition", in *ACL*, 2019, pp. 1430-1440.