علوم رایانش و فناوری اطلاعات
نشریه علمی انجمن کامپیوتر ایران
مجلد ۱۷، شماره ۲، ۱۳۹۸
صفحات ۱۰۲- ۹۵
مقاله عادی

عدالت‌یار هوشمند با استفاده از شبکه‌های عصبی بازگشتی

فرناز صباحی*[۱]

*نویسنده مسئول
، دریافت: ۹۸/۱۱/۲۸، بازنگری: ۹۹/۰۲/۰۹، پذیرش: ۹۹/۰۳/۱۶

[۱] استادیار، مهندسی برق-کنترل، دانشکده فنی و مهندسی، دانشگاه ارومیه، ارومیه، ایران

چکیده

این مقاله، با بسط مفهوم تمرکز بر ویژگی‌ها، در استفاده از شبکه حافظه کوتاه- ماندگار دوطرفه (biLSTM)، یک معماری جدید برای مسائل کاربردی پیشنهاد می‌دهد. biLSTM، فرایند گذشته و آینده ویژگی‌ها را به‌طور کامل می‌تواند منعکس کند. سیستم پیشنهادی به مطالعه موردی در مسائل قضایی اعمال‌شده است. در سیستم عدالت‌یار هوشمند پیشنهادشده، برای تصمیم مؤثرتر بعد از biLSTM از دو رمزگذار استفاده‌شده است و دارای لایه‌ای مبتنی بر دانش خبرگان می‌باشد. در این روش با مشاهده اجزای پرونده ابتدا نوع مؤلفه‌ها بررسی می‌شود، سپس کلیدی بودن مؤلفه در اصلاح وزن‌ها موردتوجه قرار می‌گیرد. روش پیشنهادی در دو طرح مختلف ارائه‌شده است، در هر دو طرح ابتدا biLSTM هم بر روی مؤلفه‌های پرونده و هم بر روی حکم که دو بخش تبرئه و محکوم است اعمال می‌شود. دقت عملکرد بر اساس تمرکز بر روی مؤلفه‌های مؤثرتر مشخص می‌شود. طراحی این معماری بر اساس اشتراک‌گذاری وزن‌ها در زمان آموزش توسط رمزگذارها می‌باشد در طرح اول ابتدا مفهوم تمرکز بر ویژگی‌ها اعمال می‌شود و سپس در لایه‌های بعد هوش جمعی اعمال می‌شود. در طرح دوم هوش جمعی خبرگان در قالب یک تابع عضویت فازی به آن اعمال می‌شود. نتایج سیستم مشاور پیشنهادی در مورد مطالعاتی قضایی با روش‌های دیگر مقایسه شده‌اند که برتری روش پیشنهادی مشخص‌شده است. روش پیشنهادی با طراحی یک الگوی مناسب و به‌کارگیری اکثر عامل‌های دخیل و شناخت تأثیرگذاری آن‌ها درگرفتن یک تصمیم درست‌تر در زمان کوتاه‌تر می‌تواند بسیار کمک‌کننده باشد و متعاقباً هزینه‌های تشکیل دادگاه‌های تجدیدنظر و اطاله دادرسی را کاهش می‌دهد و حس اعتماد جامعه به سیستم قضا را افزایش می‌دهد.

کلمات کلیدی: تصمیم‌گیری، پرونده‌های قضایی، LSTM دوطرفه

۱- مقدمه

ما در جهانی زندگی می‌کنیم که ازلحاظ در دسترس بودن اطلاعات و میزان عدم قطعیت رشد فزاینده‌ای دارد. این رشد فزاینده در مسائل قضایی نیز وجود دارد. به‌طورکلی، عدالت قضایی یک فرایند است که با تحقیقات اولیه شروع می‌گردد و با استفاده از اطلاعات که معمولاً با زبان محاوره‌ای بیان می‌شود مراحل را طی می‌کند و درنهایت با صدور رأی برای متهم خاتمه می‌پذیرد. دانستن قوانین و استفاده از آن‌ها فقط بخش کوچکی از پیچیدگی فرآیند قضاوت است. قضاوت با عدم قطعیت زیادی مواجه است. درواقع، تعیین گناهکاری یا بی‌گناهی متهم به عامل‌های زیادی بستگی دارد. رابطهٔ علیت در جرم‌شناسی که منجر به قضاوت می‌شود دایره‌ای است و یک عامل ممکن است منجر به نتیجه‌ی واحد نشود و نتایج متعدد و حتی متضاد داشته باشد[۱].

پیچیدگی داوری به عوامل متعددی ارتباط دارد ازجمله اندازه دادگاه (یک قاضی و یا جمعی از قضات)، محتوای قوانین، شرایط پرونده‌ها، استانداردهای موجود. از طرف دیگر، حکم‌های قاضی نیز باید چندین ویژگی داشته باشند [۲]: آن‌ها باید «مبتنی بر قانون، یقین کافی و در جهت حمایت از فعالیت‌های ارزشمند باشند.» درواقع، تصمیم‌گیری قضایی مبتنی بر مجموعه‌ای از عوامل ازجمله: تجزیه‌وتحلیل قوانین، پرونده‌های مشابه قبلی، روش‌های تجربی و عقل سلیم می‌باشد که به شخصیت شخصی قاضی، اولویت‌های ایشان، ارزش‌ها، پس‌زمینه و تحصیلات آن‌ها بستگی دارد. به‌طور مشخص، قضات که در معرض گفته‌های احتمالاً غیرقابل استناد، ارزیابی مستندات گهگاه مشکوک، تجزیه‌وتحلیل قوانین احتمالاً کمتر مرتبط، شناسایی قرائن احتمالاً پنهان در شرایطی که شاید برای انحراف توجه قضات از حقیقت، ایجاد می‌شود و در حضور تلاش‌هایی که معمولاً درصدد تحریک احساسات قضات دارند، تصمیم‌گیری کنند و تصمیم درست را بگیرند.

اطلاعات مربوط به پرونده قضایی به‌صورت مفاهیم ادراکی در شکل جمله در زبان محاوره‌ای بیان‌شده‌اند می‌باشند. ادبیات نظری درباره تصمیم‌گیری قضات به دو بخش عمده تقسیم می‌شود: نظریه انتخاب منطقی [۳] و نظریه‌هایی که بر اساس قوانین و احساس شهودی [۴] هستند. اما به علت اینکه نظریه اول بر اساس مدل

است یک ناسازگاری در عمل به وجود می‌آید. اما نظریه دوم که بر اساس نظر متخصصان و خبرگان است بیشتر بر مبنای واقعیت می‌باشد. درواقع، اولین دید هوش مصنوعی به عدالت قضایی و قانون به استفاده ازنظر کارشناسان مربوط می‌شود. بازیابی اطلاعات آماری از پایگاه داده برای کمک به قضات در مقالات زیادی به‌کاربرده شده است که می‌توان به پرونده‌های کیفری کانادا در [۵] اشاره نمود. نویسندگان در [۶]، یک استدلال مبتنی بر پرونده‌های مشابه دادرسی برای سازمان‌بندی دفاعیات و ارزیابی دادرسی‌های قبلی ارائه داده‌اند. از دیگر روش‌های مطرح می‌توان به استفاده از سیستم پشتیبان تصمیم‌گیری [۷]، هستی‌شناسی [۸, ۹]، شبکه‌های بیزین [۱۰]، معناشناسی وب [۱۱] و داده‌کاوی [۱۲] اشاره نمود. در [۱۳]، استفاده از شبکه عصبی برای دادن حکم در مواردی که قانون واضح نیست با استفاده از موارد دادرسی قبلی آدرس‌دهی شده است که در ورودی شبکه عصبی آن از بیست‌وهفت نوع از جرم‌های اتفاق افتاده استفاده‌شده است و خروجی آن به سه دسته کلی: اعدام-حبس ابد و جریمه نقدی با ۱۰ سال حبس و جریمه نقدی تقسیم‌شده است. همچنین، در [۱۴] از درخت‌های تصمیم برای تشخیص اجزای جرم استفاده‌شده است. تشخیص قوانین مربوط به جرم بر اساس یک جدول جستجو است که با تحلیل بازگشتی فازی محکومیت مجرم تعیین می‌شود. روشی مبتنی بر سیستم‌های چندعاملی برای استدلال سلسله مراتبی در حوزه تصمیم‌گیری‌های قضایی در [۱۵] پیشنهادشده است. همچنین استفاده از نظریه دمپستر-شافر برای ارزیابی مستندات در [۱۶] ارائه‌شده است که به ارزیابی میزان خطر قبول مستندات پرداخته‌شده است. نویسندگان در [۱۷] از منطق توصیفی شهودی برای آنالیز موضوع تضاد قوانین استفاده کرده‌اند. اکثر روش‌های بالا اگرچه به نتایجی مستدل رسیده‌اند، اما فقط طیف محدودی از عدم قطعیت‌ها را در نظر گرفتند. استفاده از مفهوم اعتبار علاوه بر مفهوم احتمال و فازی در قضاوت پرونده‌های کیفری نیز پیشنهادشده است[۱۸, ۱۹].

در سال‌های اخیر استفاده از روش‌های یادگیری عمیق در پیش‌بینی‌ها و تصمیم‌گیری‌ها افزایش پیداکرده است [۲۰, ۲۱]. یادگیری عمیق شاخه‌ای از هوش مصنوعی است که شامل دسته‌ای از الگوریتم‌هایی است که مفاهیم سطح بالا و انتزاعی (مفهومی) را در لایه‌های مختلف مدل می‌کند. در مقاله [۲۲]، نویسندگان برای تحلیل سیگنال‌های مغزی از حافظه کوتاه‌مدت طولانی یک‌طرفه (LSTM) در سری‌های زمانی استفاده کردند. در تشخیص خطای یک مبدل چند سطحی از LSTM در [۲۳] استفاده‌شده است. استفاده از یادگیری عمیق در تصمیم‌گیری قضایی در [۲۴] گزارش‌شده است.

نکته قابل‌توجه این است که در تصمیم‌گیری قضایی دودسته اطلاعات وجود دارد، اولی نظر متخصصان است که عدم قطعیتی به شکل خطا یا بی‌خبری و ابهام دارد و دومی داده‌های آماری است که در اغلب موارد ناکافی‌اند. به‌طور مشخص‌تر، فرایند دادگاه بیشتر بر اساس متقاعد کردن [۲۵] مصطلح شده است؛ بنابراین ممکن است همیشه نتوان به همان نتایج اول رسید حتی اگر شرایط عیناً تکرار شده باشد. باوراندن شواهد، منجر به اثبات ادعا می‌شود، اما باور شواهد شامل متقاعد کردن است که تا حد زیادی متکی بر درک شهودی است. درواقع، متقاعد کردن متکی بر ترکیبی از احساس، تجربه، حس مشترک، پیش‌داوری و تمایل یا حتی سلیقه است. در دادگاه، استدلال ممکن است شامل مفاهیم غیرقابل‌اعتماد و یا نامربوط باشد، اما هنوز هم تا حدی قانع‌کننده باقی بماند. علاوه بر این، زمانی که قانون نامشخص است، از قضات انتظار می‌رود شرایط را تفسیر کنند و یا روش وحدت رویه را پیش بگیرند و حکم‌هایی با توجه کمتر به قوانین ایجاد کنند. ازاین‌رو، با نقاط حقوقی متمایز و دیدگاه‌های فردی به عدالت، قضات مختلف ممکن است به رأی‌های مختلف در مورد یک پرونده قضایی یکسانی برسند. این نشان می‌دهد که اعتبار یک تصمیم قاضی باید تا حد امکان کمتر به قلمرو برداشت فردی از مسائل قانونی مربوط شود. به عبارت روشن‌تر، در قضاوت، گرایش‌های قاضی و شخصیت او در تصمیماتی که می‌گیرد نقش دارد. داده‌های یادگیری به‌طور مستقیم وجود ندارند و مشاهده در قضاوت، خاص خود قاضی است. بنابراین، ازآنجاکه انتخاب قاضی به عدم قطعیت‌های زیادی وابسته است و انتخاب نهایی قاضی را نمی‌توان همیشه کاملاً تائید یا رد کرد، برای یک تصمیم قضایی درست باید عامل‌های مهم و مختلفی را مدنظر داشت. علاوه بر این، باید تأثیر این عوامل در پرونده‌های گذشته نیز بررسی شوند و از آن‌ها در پرونده‌های آینده استفاده شود.

در این مقاله، روش پیشنهادی استفاده از ساختار دوطرفه حافظه biLSTM می‌باشد که این biLSTM در قالب استخراج و استنتاج برای داده حکم استفاده می‌شود. برای مقابله با عدم قطعیت موجود در مسئله قضاوت از مفهوم تمرکز[۱] استفاده کردیم که با هوش جمعی که توسط یک پرسشنامه از یک سری از متخصصان حقوق و قضا فراهم‌شده است بهبود داده‌شده است

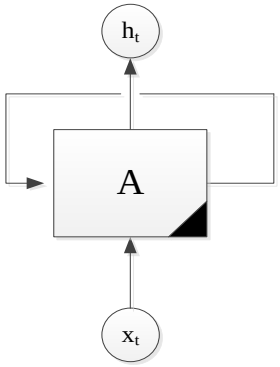
لازم به ذکر است که روش پیشنهادی نقش یک مشاور را دارد. اگر متخصصان در دسترس باشند سیستم ما در پرونده‌ها رأیی را پیشنهاد می‌دهد که به دلیل استفاده کردن از هر دو دانش آماری و دانش خبرگان درجه اطمینان بالایی دارد پس متخصصان می‌توانند به‌عنوان یک مشاور مطمئن به آن مراجعه کنند. درواقع، ترکیب دانش آماری با دانش خبرگان به‌منزله ترکیب منابع مربوط به مفهوم احتمال و امکان[۲] است. مفهوم آمار و احتمال بر مبنای داده‌های آماری تعریف می‌شوند. دانش خبرگان بیشتر بر اساس زبان طبیعی است که قابلیت منطق فازی مدل کردن این نوع دانش است. این دو مفهوم بر مبنای دو دسته اطلاعات متفاوت‌اند که جنبه‌های متفاوت دانش را بیان می‌کنند پس ترکیب این دو می‌تواند منجر به نتیجه بهتر شود [۲۶]. بنابراین، در مسئله خاص قضاوت، قاضی می‌تواند به سیستم پیشنهادی همچون یک مشاور توجه کند تا رأی نهایی او با اطمینان بیشتری صادر شود. در صورت نبود متخصصان با توجه به نحوه ساختار سیستم پیشنهادی، رأی ارائه شده توسط سیستم درجه اطمینان بالاتری نسبت به روش‌های موجود دارد.

ساختار این مقاله به‌صورت زیر است: در بخش بعدی، به‌طور خلاصه به شبکه‌های عصبی بازگشتی ازجمله حافظه کوتاه-ماندگار یک‌طرفه و دوطرفه پرداخته‌شده است. با ارزیابی مؤلفه‌های مؤثر در پرونده‌های قضایی به معرفی و توضیح معماری عدالت‌یار هوشمند پیشنهادی در بخش سوم پرداخته‌شده است. نتایج شبیه‌سازی در بخش چهارم و نتیجه‌گیری در بخش پنجم آمده است.

۲- شبکه‌های عصبی بازگشتی

شبکه‌های عصبی بازگشتی استاندارد یک حلقه بازگشتی دارند که باعث می‌شود تا اطلاعاتی که از لحظات پیش‌تر به دست آورده شده بودند حذف نشوند زیرا در آینده ممکن است به‌کاربرده شوند. یک‌شکل کلی از این نوع شبکه‌های عصبی در شکل ۱ نشان داده‌شده است. همان‌طور که در این شکل دیده می‌شود x_t به‌عنوان ورودی وارد بخشA می‌شود و مقدار h_t را در خروجی ایجاد می‌کند. وجود حلقه در شبکه عصبی بازگشتی استاندارد باعث می‌شود که اطلاعات از یک مرحله‌به‌مرحله بعدی ارسال گردد.

اگر شبکه عصبی بازگشتی استاندارد باز شود مانند یک زنجیر خواهد بود؛ این شبکه‌ها مربوط به‌توالی و فهرست‌ها می‌باشند. در حقیقت، شبکه‌های عصبی بازگشتی استاندارد مناسب‌ترین انتخاب برای کار با داده‌های متوالی هستند. شبکه‌های عصبی بازگشتی موفقیت‌های زیادی در کاربردهای عملی داشته‌اند. اگر فقط نیاز به اطلاعات مربوط به گذشته نزدیک برای بررسی مسئله باشد، شبکه‌های عصبی بازگشتی استاندارد بهترین گزینه برای این کار می‌باشند اما اگر فاصله بین اطلاعات گذشته موردنیاز و جایی که به این اطلاعات نیاز است زیاد باشد (اصطلاحاً به این موقعیت وابستگی زمانی بلندمدت می‌گویند)، شبکه‌های عصبی بازگشتی استاندارد قدرتشان را در یادگیری این اطلاعات از دست می‌دهند [۲۷].



شکل ۱- شبکه عصبی بازگشتی

این مشکل توسط شبکه‌های عصبی بازگشتی حافظه کوتاه-ماندگار که به‌اختصار LSTM نامیده می‌شوند حل‌شده است [۲۸]. هر واحد LSTM همان‌طور که در شکل (۲) دیده می‌شود، برعکس شبکه عصبی بازگشتی استاندارد که فقط جمع سیگنال‌های ورودی را از یک تابع فعال‌سازی عبور می‌دهد، از یک حافظه در زمان بهره می‌برد [۲۹]. درواقع، شبکه‌های LSTM مانند شبکه‌های عصبی بازگشتی استاندارد ساختار زنجیره مانند دارند ولی بخش تکرارشونده آن‌ها به‌جای یک‌لایه، چهار لایه دارد که به‌طور ویژه‌ای با یکدیگر در تعامل هستند. بخش اصلی LSTM سلول حالت[۳] است. LSTM می‌تواند با استفاده از دروازه‌ها[۴] اطلاعات جدیدی را به سلول حالت اضافه یا حذف کند. دروازه‌ها که از یک‌لایه سیگموید[۵] و یک عملگر ضرب نقطه‌ای ساخته‌شده‌اند راهی برای ورود اختیاری اطلاعات هستند.

فعال‌سازی واحد LSTM به‌صورت زیر می‌باشد:

$$h_t = \Gamma_0 tanh(C_t) \quad (۱)$$

که در آن Γ_0 دروازه خروجی و C_t سلول حافظه است. Γ_0 میزان محتوای حافظه را کنترل می‌کند و به طریق عبارت زیر محاسبه می‌شود:

$$\Gamma_0 = \sigma(W_0.[h_{t-1}, X] + b_0) \quad (۲)$$

که در آن σ تابع فعال‌سازی سیگموید و W_0 ماتریس وزن است. بخشی از حافظه که باید حذف شود توسط Γ_f و بخشی از محتوای جدید که باید به حافظه اضافه شود توسط Γ_u به‌صورت‌های زیر کنترل می‌شود:

$$\Gamma_f = \sigma(W_f.[h_{t-1}, X] + b_f) \quad (۳)$$
$$\Gamma_u = \sigma(W_u.[h_{t-1}, X] + b_u)$$

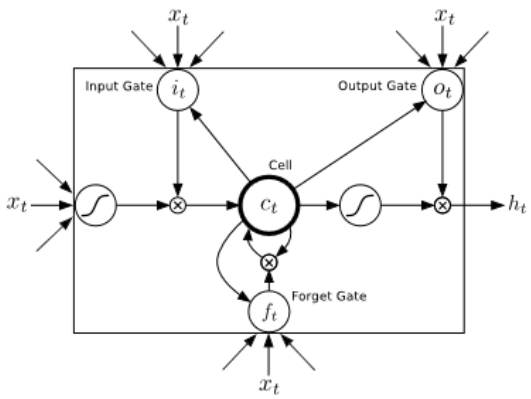
که W_f ماتریس وزنی است که رفتار بردار Γ_f را کنترل می‌کند که مقداری بین صفر و یک دارد. این بردار در ادامه در C_{t-1} ضرب خواهد شد. اگر $\Gamma_f = 0$ باشد محتوای C_{t-1} حذف می‌شود. اگر $\Gamma_f = 1$ بشد محتوای C_{t-1} حفظ می‌شود. مقادیر $0 < \Gamma_f < 1$ نیز مبین این است که بخشی از محتوای C_{t-1} حذف‌شده و بخشی از محتوای C_{t-1} حفظ شده است.

سلول حافظه با فراموشی نسبی حافظه فعلی و اضافه کردن محتوای جدید به‌صورت $\hat{C}_t$ به‌صورت زیر به‌روزرسانی می‌شود:

$$C_t = \Gamma_f . C_{t-1} + \Gamma_u . \hat{C}_t \quad (۴)$$
$$\hat{C}_t = tanh(W_c.[h_{t-1}, X] + b_c)$$

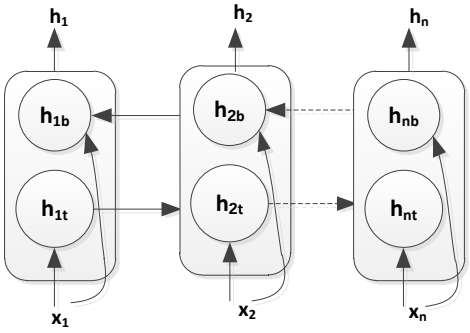
در انتها نیز برای اینکه مشخص کنیم در خروجی از چه محتوایی باید استفاده کنیم از دروازه خروجی بهره می‌بریم. شیوه کار به‌صورت زیر است:

$$h_t = \Gamma_0 tanh(C_t) \quad (۵)$$



شکل ۲- معماری LSTM [۳۰]

biLSTM یا LSTM دوطرفه نیز یکی از روش‌های یادگیری عمیق است[۳۱]. همان‌طوری که در شکل (۳) دیده می‌شود LSTM دوطرفه، به این صورت است که در زمان آموزش، برای بررسی داده‌ها یک‌بار از ابتدا به انتها می‌رویم و وزن‌ها و مقادیر بایاس اصلاح می‌شوند، یک‌بار دیگر از انتها به ابتدا فرآیند آموزش انجام می‌شود و وزن‌ها و مقادیر بایاس تنظیم می‌شوند و در انتها این دو، در کنار هم در خروجی قرار می‌گیرند.



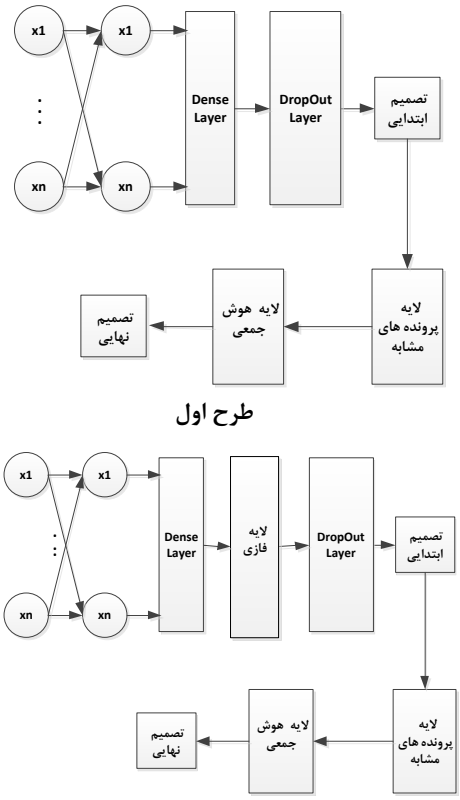
شکل ۳- معماری biLSTM

۳- روش پیشنهادی: عدالت‌یار هوشمند

در این بخش به جزئیات روش پیشنهادی با عنوان عدالت‌یار هوشمند می‌پردازیم. در ابتدا، با مشاهده اجزای پرونده قضایی، نوع مؤلفه‌ها بررسی می‌شود و کلیدی بودن مؤلفه در اصلاح وزن‌ها موردتوجه قرار می‌گیرد. به‌طور مثال اگر "مؤلفه ۱" نسبت به "مؤلفه ۲" اهمیت بیشتری داشت، وزن مؤلفه ۲ به میزان کمی اصلاح می‌شود. در ابتدا برای تعیین مؤلفه‌های کلیدی و میزان اهمیت آن‌ها یک مطالعه موردی انجام می‌شود. این مطالعه به‌وسیله یک پرسشنامه به‌عنوان ابزار جمع‌آوری اطلاعات بر روی جمعی از متخصصان قضاوت و وکلا و دانشجویان در حال تحصیل حقوق جزا در شهر مشهد انجام‌شده است. در فاز اول مطالعه و برای طراحی پرسشنامه به شناسایی مؤلفه‌های مؤثر پرداخته‌شده است. برای این کار از منابع مختلف مانند کتب قضا، پرونده‌های واقعی و متخصصان استفاده‌شده است. مؤلفه‌های تعیین‌شده بر اساس تبادل‌نظر با متخصصین درنهایت به دوازده مورد رسید که عبارت‌اند از: دروغ، اقرار متهم، اقرار همدست متهم، مدارک علمی، مدارک فیزیکی، مدارک دیجیتالی، نظر متخصص، اسناد، مدارک توصیفی، گواهی شاهدان، تکرار همان جرم درگذشته و سوءپیشینه.

در پرسشنامه از پاسخگویان خواسته شد که نظرات خود را نه به‌صورت یک عدد بلکه به‌صورت بازه بیان کنند که حد پایین این بازه معرف کم‌ترین اهمیت مؤلفه و حد بالا معرف بیشترین اهمیت آن در محکومیت یا تبرئه است. محدوده این بازه فاصله عددی بین صفر و ده بود. مخاطبین پرسشنامه عبارت بودند از: دانشجویان تحصیلات تکمیلی سال آخر در مقطع ارشد و اساتید دانشگاه‌ها به تعداد حدود نوزده

نفر، متخصصان وکالت و قضاوت شاغل در شهر مشهد حدود دوازده نفر و حدود چهار نفر از متخصصین در سایر رشته‌های قضایی که همگی شاغل به تحصیل بودند. پرسشنامه توسط مجری توزیع‌کننده برای هرکدام از پاسخگویان در هنگام تکمیل توضیحات لازم ارائه می‌شده است ضمن اینکه به سؤالات آن‌ها پاسخ داده می‌شده است. سپس کلیه مؤلفه‌ها در فایل اکسل آورده شد. مقادیر مؤلفه‌ها و میانگین و انحراف معیار آن‌ها مقایسه و در یک فایل ذخیره شد. در بررسی پایای درونی پرسشنامه ضریب آلفای کرونباخ [۳۲] پس از ورود اطلاعات به کامپیوتر با نرم‌افزار SPSS برابر ۹۰/۴ درصد برآورد شد که نشان‌دهنده ثبات بالای پرسشنامه می‌باشد. معماری روش عدالت‌یار هوشمند در دو طرح پیشنهادی در شکل (۴) آورده شده است. در شکل، x_i ها معرف دنباله کلمات جملات پرونده و حکم می‌باشند.



شکل ۴- ساختار کلی معماری پیشنهادشده

در روش پیشنهادی، ابتدا biLSTM هم بر روی پرونده که ورودی است و هم بر روی حکم که دو بخش تبرئه و محکوم است اعمال می‌شود. در بخش اول کار، سیستم خلاصه‌سازی پرونده و کشف مؤلفه‌ها و حکم از پرونده‌ها را انجام می‌دهد (شکل ۵). دقت عملکرد بر اساس تمرکز بر روی مؤلفه‌های مؤثرتر مشخص می‌شود. مؤثر بودن مؤلفه به همان ترتیبی تعیین می‌شود که در همین بخش توضیح داده‌شده است. طراحی این معماری بر اساس اشتراک‌گذاری وزن‌ها در زمان آموزش توسط رمزگذارها می‌باشد. سپس دو بخش با یکدیگر ترکیب می‌شوند. برای دریافت جمله‌ها به‌عنوان ورودی دو استراتژی معکوس کردن ورودی و تکرار دوباره استفاده‌شده است. این استراتژی‌های برگرفته از مراجع مختلف می‌باشند، که باعث افزایش دقت در تصمیم‌گیری نهایی می‌شوند. معکوس کردن ورودی یعنی ترتیب کلمات از آخر به ابتدا و با همان حکم وارد شوند. این کار یک کار خنثی است فقط دقت را بالا می‌برد این استراتژی برای با اول در[۳۳] معرفی‌شده است. پردازش با تکرار دوباره ورودی با همان حکم نیز انجام می‌شود. در این حالت ما ورودی را دو بار تکرار می‌کنیم یعنی اگر ورودی "الف" باشد حالا ورودی" الف ، الف" می‌شود. از دیدگاه احتمالاتی این کار بی‌معنی است زیرا به‌جای یادگیری $p(y|x)$ سعی در یادگیری $p(y|x,x)$ داریم اما نشان داده‌شده است که این کار باعث افزایش قابل‌توجه دقت نتایج می‌شود. چون پردازش دوباره ورودی بدون به دست آوردن نتایج باعث می‌شود که biLSTM فرصت برطرف کردن اشتباه احتمالی را داشته باشد[۳۴]. بعد از انتخاب ورودی، این جمله‌ها با کمک وسیله‌ای به نام Tokenizer لغت به لغت جداشده و با مشابهت‌سازی با مؤلفه‌های مؤثر به آرایه‌ای از مؤلفه‌ها تبدیل می‌شوند. برای یکسان‌سازی تمام پرونده‌ها، این آرایه داده، به‌اندازه بیشترین مؤلفه ممکن در یک پرونده که برابر با ۱۲ است، فضا اختصاص داده می‌شود. این مؤلفه‌ها با توجه به اینکه چند بار در کل پرونده در موارد مختلف تکرار شده‌اند، رتبه‌بندی می‌شوند. در زمان آموزش، نگاشت بین این اعداد و مؤلفه‌ها در آرایه دیگری ذخیره می‌شود. سپس این آرایه از مؤلفه‌های کلیدی به‌منظور انجام عمل یادگیری به‌صورت یک آرایه شامل اعداد تصادفی ۳۰۰ تایی بین ۰ تا ۱ در فضای برداری ذخیره می‌شود که نقش تعبیه ساز کلمه را به عهده دارد. عدد ۳۰۰ بر اساس سعی و خطا و با توجه به میزان دقت اجرا به‌اندازه بردار تعیین‌شده است. لازم به ذکر است که بررسی اندازه بردار به میزان دقت در تحقیقی که در مرجع [۳۵] آورده شده است انجام‌شده است بعد از به دست آمدن آرایه ۱۲ در ۳۰۰ تایی، این لایه به لایه biLSTM متصل می‌شود که خروجی این لایه ۶۰۰ تایی است.

سپس، برای ایجاد دقت بیشتر در تعیین حکم از رمزگذاری تجمع متوسط[۶] با اندازه پنجره ۲ بعد از لایه biLSTM استفاده‌شده است (R_{avg})، در مرحله اول از این رمزگذار برای تولید بردار اصلی مربوط به پرونده استفاده می‌شود. سپس با به دست آوردن وزن مؤلفه‌های مؤثر با استفاده از ارجاع به خبرگان که در مرحله دانش مؤلفه‌های مؤثرتر تولیدشده است، استفاده می‌شود. ورودی این لایه ۶۰۰ بعد و خروجی این لایه ۳۰۰ بعد دارد. لایه R_{avg} که یکی بالای مؤلفه‌ها و یکی بالای حکم است، که خروجی‌های آن‌ها کنار هم ۶۰۰ بعد می‌شود که با خروجیbiLSTM که ۶۰۰ بعد است جمع می‌شود. سپس، تحت تابع فعال‌سازی تانژانت هیبرید قرار می‌گیرد.

$$M = tanh(W^y Y + W^h R_{ave} \otimes e_L) \qquad (۶)$$

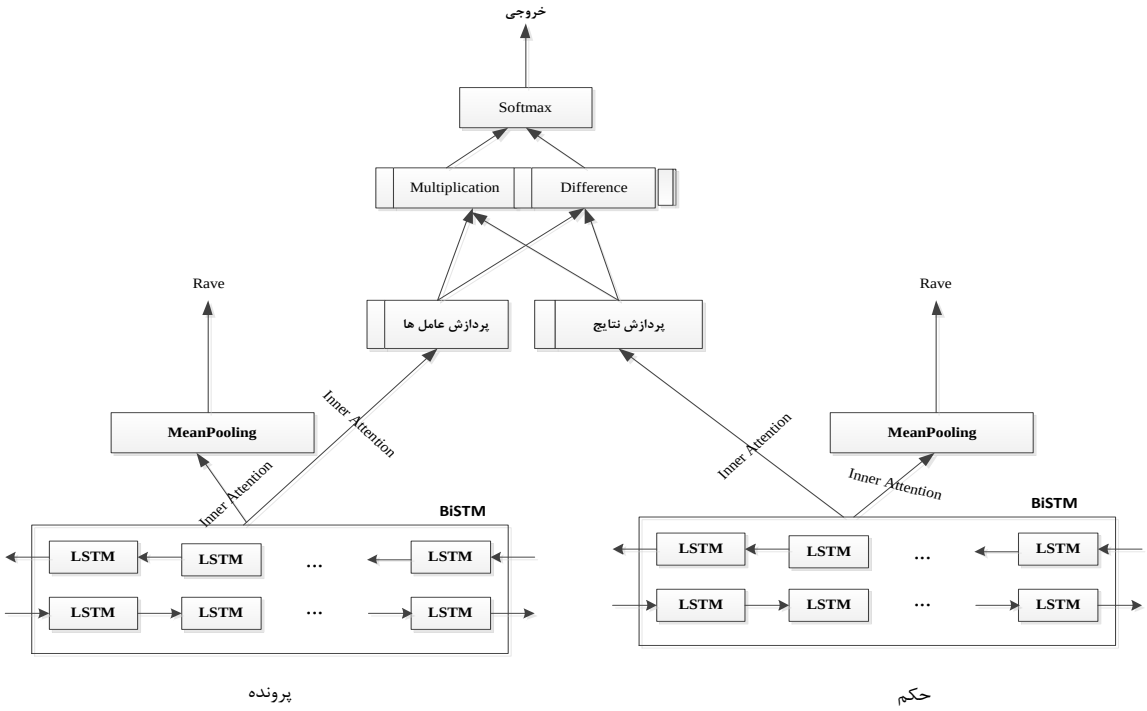
در این قسمت، نیز دو بخش جداگانه یکی برای مؤلفه‌ها و یکی برای حکم مورداستفاده قرار می‌گیرد. سپس تابع فعال‌سازی بیشینه هموار[۷] قرار می‌گیرد، ورودی ۶۰۰ بعد و خروجی ۶۰۰ بعد است این ضرب باعث می‌شود تأثیر شرایط برای آموزش تغییر کند، این لایه در خروجی ۶۰۰ بعد دارد.

$$\alpha = softmax(W^T M) \qquad (۷)$$
$$\mathrm{R}_{att} = \alpha Y^T$$

در زمان آموزش، ۳ بردار وزن W^hو W^Tو W^y نیز آموزش داده می‌شوند تا به‌درستی تأثیر شرایط مختلف را نشان دهند. Y یک ماتریس است که شامل خروجی biLSTM است، α مبین بردار توجه می‌باشد. وR_{att} مبین توجه وزن‌دار شده می‌باشد.

برای استخراج رابطه بین پرونده و حکم سه عمل زیر را انجام می‌دهیم:

۱. قرار دادن هر دو بخش به‌صورت لغت کنار هم (۱۲ مؤلفه و یک حکم)
۲. ضرب دولایه ۶۰۰ بعدی بردار پرونده و حکم که یک خروجی ۶۰۰ بعد ایجاد می‌کند.
۳. محاسبه تفاوت دولایه ۶۰۰ بعدی مربوط به پرونده و حکم که یک خروجی ۶۰۰ بعد ایجاد می‌کند.



شکل ۵- جزئیات مقدماتی معماری پیشنهادشده

این بخش‌های ترکیب برگرفته از مرجع [۳۶] می‌باشد. با کنار هم قرار دادن این سه بخش بالا یک خروجی ۱۲۱۳ بعدی خواهیم داشت. درنهایت تابع فعال‌سازی بیشینه هموار قرار دارد. خروجی واقعی برای حکم محکوم به‌صورت $\begin{bmatrix}1\\0\end{bmatrix}$ و برای حکم تبرئه به‌صورت $\begin{bmatrix}0\\1\end{bmatrix}$ در نظر گرفته شده است.

برای آموزش و به‌روز کردن بردارهای وزن از روش پس انتشار خطا بر مبنای زمان[۸] استفاده می‌شود. پس انتشار خطا، برای به دست آوردن و به‌روزرسانی ماتریس‌های وزن مورداستفاده می‌شود. در انتشار روبه‌جلو در biLSTM و برای به دست آوردن سلول‌های فعلی C و حالت‌های پنهان h ورودی به لایه‌های پنهان تزریق می‌شود و خروجی در واحد زمان به دست می‌آید. تابع هزینه براساس تفاوت خروجی به‌دست‌آمده و خروجی واقعی به دست می‌آید و در مسیر برگشت باعث به‌روزرسانی ماتریس وزن می‌شود. تابع هزینه بخش آموزش بر اساس تابع خطای آنتروپی متقاطع[۹] می‌باشد که با تعریف y همچون خروجی واقعی و z همچون خروجی شبکه به‌صورت زیر تعریف می‌شود:

$$L(z,y,t) = -[y\log(z) + (1-y)\log(1-z)] \qquad (۸)$$

بنابراین، تابع هزینه کلی بین دو زمان به‌صورت زیر تعریف می‌شود:

$$L_{total}(t_0, t_1) = \sum_{t=t_0}^{t_1} L(z,y,t) \qquad (۹)$$

گرادیان‌ها برای به دست آوردن تغییرات وزن به‌صورت زیر می‌باشند:

$$\Delta w_{ij} = -\eta \frac{\partial\ L_{total}(t_0, t_1)}{\partial w_{ij}} \qquad (۱۰)$$

که w_{ij} مربوط به بردارهای وزن W^h و W^T و W^y است. همچنین، η نرخ یادگیری می‌باشد.

درروش پیشنهادی از داده‌های ورودی ابتدا به دو بخش: بخش آموزش و آزمون تقسیم می‌شوند سپس بخش آموزش برای آموزش biLSTM به کار می‌روند و بخش آزمون به‌صورت آنلاین به کار می‌رود. حالا با یک آرایه حداکثر ۱۲تایی مؤلفه‌ها و حکم سروکار پیدا می‌کنیم که وارد لایه دوم در شکل ۴ می‌شود. درواقع، خروجی‌های لایه اول، ورودی‌های لایه Dense برای تجمع متوسط و سپس لایه Dropout و با استفاده از بیشینه هموار برای گرفتن حکم اولیه استفاده می‌کند.

بعد از تولید ابتدایی تصمیم، این تصمیم در لایه پرونده‌های مشابه در قالب یک وزن که به‌صورت ضریب عمل می‌کند بهبود پیدا می‌کند. لایه پرونده‌های مشابه، بر اساس داده‌های آماری در دسترس، عمل می‌کند. داده‌های آماری در دسترس در جدول ۱ آمده است.

در اینجا دو طرح از روش پیشنهادی معرفی می‌شود:

طرح اول: قبل از تعیین تصمیم نهایی، تصمیم اولیه تحت هوش جمعی قرار می‌گیرد. سپس، عدد نهایی مربوط به‌حکم تبرئه و محکوم باهم مقایسه می‌شود. هرکدام بزرگ‌تر بود حکم را تعیین می‌کند.

طرح دوم: تعیین تصمیم نهایی با استفاده از تصمیم فازی بر مبنای هوش جمعی انجام می‌شود. برای این کار با در نظر گرفتن تصمیم اولیه با $d_{Initial}$ و تصمیم نهایی با d_{Final} یک تابع عضویت خطی به‌صورت زیر در نظر می‌گیریم:

$$\mu(d_{Final}) = \begin{cases} 0 & d_{Initial} < 0 \\ d_{Final} & 0 < d_{Initial} < 1 \\ 1 & d_{Initial} > 1 \end{cases} \qquad (۱۱)$$

سپس،

$$\mathrm{R}_{attention} = \mathrm{Y}\alpha^T \times \mu(d_{Final}) \qquad (۱۲)$$

ازآنجاکه برای هر بخش حکم بخش گناهکار و بخش تبرئه این عمل انجام می‌شود حالا با مقایسه دو عدد به‌دست‌آمده حکم تعیین می‌شود.

لایه هوش جمعی رأی را بر اساس نظر جمعی از قضات و متخصصان بهبود می‌دهد. با این لایه، رأی صادرشده مبنی بر محکومیت یا تبرئه از تجربه سایر خبرگان نیز بهره می‌برد. بر این اساس همان‌طور که در بخش نتایج خواهیم دید این بهبود در اکثر مواقع باعث می‌شود که به دادگاه تجدیدنظر نیاز پیدا نشود و یا در صورت تشکیل این دادگاه رأی تغییری نکند.

جدول ۱- داده‌های آماری در دسترس

تعداد کل تبرئه اتهام	۱۴	تعداد کل اثبات اتهام	۱۵۳
اقرار متهم	۱	اقرار متهم	۴۷
دروغ	۰	دروغ	۲
اقرار همدست متهم	۲	اقرار همدست متهم	۴
مدارک علمی	۲	مدارک علمی	۲۲
مدارک فیزیکی	۳	مدارک فیزیکی	۲۳
مدارک دیجیتالی	۰	مدارک دیجیتالی	۲
نظر متخصص	۳	نظر متخصص	۶۱
اسناد	۱	اسناد	۲
مدارک توصیفی	۰	مدارک توصیفی	۲
تکرار جرم درگذشته	۰	تکرار جرم درگذشته	۷
سوءپیشینه	۲	سوءپیشینه	۶۱
شاهد	۰	شاهد	۱۰

۴- نتایج شبیه‌سازی

در اینجا، پرونده واقعی زیر را به‌منظور ارزیابی کارایی سیستم تصمیم‌گیرنده بررسی می‌کنیم. این پرونده در عین سادگی ظاهری آن پیچیده می‌باشد، طوری که رأی نهایی آن در دادگاه تجدیدنظر مشخص‌شده است.

پرونده: در پی شکایت فردی به نام R از سرقت اموالش و مشکوک بودن او به فردی به نام N، N دستگیرشده و اقرار می‌کند که اموال مسروقه در خانه دایی‌اش Z می‌باشد. مأمورین به خانه Z می‌روند و اموال مسروقه را در آنجا کشف و ضبط می‌کنند. اظهارات Z و N به ترتیب زیر می‌باشد: N اعلام می‌کند که به همراه دایی دیگرش به سرقت می‌روند و پس از دزدیدن گاوصندوق به خانه Z می‌آیند و از او می‌خواهند که با ماشین خود بیاید گاوصندوق را بار بزند و به خانه خود ببرد. Z همه موارد بالا را تصدیق می‌کند و اظهار پشیمانی می‌کند.

در این پرونده متهمZ است و موارد زیر وجود دارد:

۱) اثر متهمان ۲) عدم سوءپیشینه ۳) متهمان دیگر ۴) مورد دزدی گاوصندوق است که کشف‌شده است ۵) متهم اقرار کرده است. اتهامات برای Z بدین قرار است: دزدی و پنهان کردن اموال دزدی. قصد داریم در هر حالت تعیین کنیم که آیا او گناهکار است یا نیست.

بعد از محاسبه تصمیم اولیه، هوش جمعی متناظر با اعتراف، شواهد فیزیکی و اعتراف شریک متهم، برای رأی محکومیت و نداشتن تکرار جرم و سوءپیشینه برای رأی تبرئه، آن‌ها را با عدد به‌دست‌آمده متناظر با لایه داده‌های آماری ترکیب می‌کنیم و وزن متناظر را نیز محاسبه می‌کنیم و درنهایت تصمیم نهایی مربوط را در هر دو طرح محاسبه می‌کنیم. آن‌وقت دیده می‌شود عدد مربوط به محکومیت برابر است با ۰/۸۹ برای طرح اول و برای طرح دوم ۰/۹۵و تبرئه برابر است با ۰/۱۱ برای طرح اول و برای طرح دوم ۰/۰۵ می‌رسیم. درنتیجه، سیستم پیشنهادی در هر دو طرح با اعتبار کافی رأی به گناهکار بودن برای پنهان کردن مال دزدی می‌دهد. برای اتهام دزدی محکومیت ۰/۳۸ برای طرح اول و برای طرح دوم ۰/۱۲ و تبرئه برای ۰/۶۲ برای طرح اول و برای طرح دوم ۰/۷۸به دست می‌آید. درنتیجه سیستم پیشنهادی رأی به برائت از اتهام دزدی می‌دهد. در دادرسی واقعی، قاضی برای هردو اتهام در دادگاه ابتدایی $\mathbb{Z}$ را محکوم کرد، اما دادگاه تجدیدنظر رأی به برائت $\mathbb{Z}$ از اتهام دزدی داد. برای بررسی نظری روش پیشنهادی نتایج آن را با نتایج سیستم‌های LSTM و سیستم‌های عصبی بازگشتی استاندارد مقایسه شده است (شکل (۶)). برای یک مقایسه عادلانه، داده‌های آموزش را ۵۸ پرونده و آزمون را ۱۵ پرونده برای همه روش‌ها در نظر می‌گیریم.

برای حذف تأثیر نوع انتخاب داده‌ها در تعیین نتیجه، اجرا ده بار، بر مبنای ۱۰ fold cross validation تکرار شده است. دلیل تکرار مربوط به کمبود داده‌های در دسترس علاوه بر جلوگیری از بیش برازش می‌باشد. همانندی رأی‌های صادره واقعی با نتایج روش پیشنهادی هر دو طرح و روش‌های LSTM و شبکه‌های عصبی بازگشتی استاندارد در جدول (۲) مقایسه شده است. منظور از نامشخص این است که به علت نزدیکی اعداد به‌دست‌آمده نمی‌توان با اطمینان کامل رأی داد. نتایج با جزییات در جدول (۳) آمده است در این جدول نتایج برای راحتی کار مقایسه نرمال شده‌اند. دیده می‌شود هر دو طرح پیشنهادی در مقایسه با نظر نهایی دادگاه واقعی نسبت به دو روش دیگر نتیجه درصد صحت بیشتری دارند.

جدول ۲- مقایسه همانندی رأی روش‌های مختلف بارأی نهایی صادره

روش پیشنهادی طرح اول و دوم	LSTM	شبکه‌های عصبی بازگشتی استاندارد	نظر نهایی دادگاه
درست	نامشخص	درست	محکوم
درست	نامشخص	نامشخص*	تبرئه
درست	غلط	درست	محکوم
درست	غلط	غلط	تبرئه
درست	درست	درست	محکوم
درست	درست	درست	تبرئه
درست	درست	نامشخص	محکوم
درست	نامشخص	نامشخص	تجدید نظر محکوم
درست	درست	درست	محکوم
درست	درست	درست	محکوم

* نامشخص یعنی که به علت نزدیکی اعداد به‌دست‌آمده نمی‌توان با اطمینان کامل رأی داد

جدول ۳- مقایسه سیستم عدالت‌یار هوشمند پیشنهادی با روش‌های دیگر روی‌داده‌های یکسان

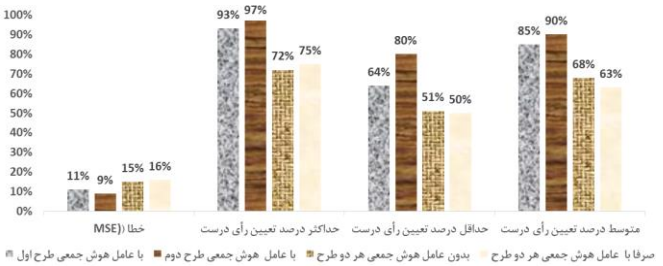
روش پیشنهادی طرح دوم		روش پیشنهادی طرح اول		LSTM		شبکه‌های عصبی بازگشتی استاندارد		نظر نهایی دادگاه
تبرئه	محکوم	تبرئه	محکوم	تبرئه	محکوم	تبرئه	محکوم	
٪۳	٪۹۷	٪۱۱	٪۸۹	٪۴۴	٪۵۶	٪۴۴	٪۵۶	محکوم
٪۸۰	٪۲۰	٪۸۲	٪۱۸	٪۵۷	٪۴۳	٪۵۳	٪۴۷	تبرئه
٪۲	٪۹۸	٪۷	٪۹۳	٪۸۴	٪۱۶	٪۳۵	٪۶۵	محکوم
٪۸۶	٪۱۴	٪۷۱	٪۲۹	٪۲۴	٪۷۶	٪۳۳	٪۶۷	تبرئه
٪۶	٪۹۴	٪۱۳	٪۸۷	۲۳٪	٪۷۷	٪۱۶	٪۸۴	محکوم
٪۸۰	٪۲۰	٪۶۴	٪۳۶	٪۷۸	٪۲۲	٪۸۷	٪۱۳	تبرئه
٪۱۶	٪۸۴	٪۱۰	٪۹۰	٪۴۰	٪۶۰	٪۴۶	٪۵۴	محکوم
٪۴	٪۹۶	٪۱۰	٪۹۰	٪۴۴	٪۵۶	٪۴۹	٪۵۱	تجدید نظر محکوم
٪۵	٪۹۵	٪۹	٪۹۱	٪۱۱	٪۸۹	٪۳۵	٪۶۵	محکوم
٪۹	٪۹۱	٪۱۱	٪۸۹	٪۴۰	٪۶۰	٪۱۶	٪۸۴	محکوم

تأثیر هوش جمعی خبرگان در جدول (۴) برای هر دو طرح معماری پیشنهادی نشان داده‌شده است. دیده می‌شود که تأثیر هوش جمعی خبرگان در طرح دوم بیشتر است. در این جدول، وقتی‌که صرفاً هوش جمعی داریم که در قالب یک سیستم فازی و پایگاه قواعد با استفاده از هوش جمعی به‌دست‌آمده آورده شده است. در این حالت، با توجه به مسئله قضاوت و جنبه تصمیم انسانی دیده می‌شود که نسبت به حالتی که جنبه یادگیری دارند افت عملکرد وجود دارد که ترکیب دانش خبرگان و داده‌های آماری در بهبود نتایج را تأیید می‌کند. این افت عملکرد نسبت به حالت

بدون هوش جمعی به دلیل مدل‌سازی مفهوم عدم قطعیت چندان زیاد نیست. مقایسه واضح‌تر در هیستوگرام در شکل ۶ نشان داده‌شده است. دیده می‌شود که نتایج بهتر طرح دوم به دلیل تأثیر هوش جمعی خبرگان می‌باشد زیرا با حذف عامل هوش جمعی خبرگان همان‌طور که در شکل۶ دیده می‌شود نتایج طرح دوم افت بیشتری نسبت به طرح اول دارد. با توجه به اینکه توابع عضویت در لایه فازی نیز بر اساس هوش جمعی به‌دست‌آمده، با حذف هوش جمعی درجه فازی یک می‌شود و این تفاوت در هر دو طرح حذف می‌شود. درواقع در طرح دوم یک‌بار قبل از تعیین رأی اولیه در غالب تابع عضویت فازی و یک‌بار بعدازآن ازنظر کارشناسان استفاده‌شده است. در طرح دوم سیستم دیرتر تصمیم می‌گیرد ولی به تصمیم بهتر می‌رسد. این طرح مناسب موارد سخت و پیچیده است و برای هر پرونده‌ای لازم نیست. اگرچه، به دلیل ورود نظر خبرگان رأی آن تطبیق بیشتری با شرایط حقیقی موجود در سیستم قضایی دارد.

جدول ۴- مقایسه تأثیر هوش جمعی در سیستم تصمیم‌یار پیشنهادی

	صرفاً عامل هوش جمعی	بدون عامل هوش جمعی	با عامل هوش جمعی	با عامل هوش جمعی
	سیستم فازی	هر دو طرح	طرح دوم	طرح اول
خطا (MSE)	٪۱۶	٪۱۵	٪۹	٪۱۱
حداکثر درصد تعیین رأی درست	٪۷۵	٪۷۲	٪۹۷	٪۹۳
حداقل درصد تعیین رأی درست	٪۵۰	٪۵۱	٪۸۰	٪۶۴
متوسط درصد تعیین رأی درست	٪۶۳	٪۶۸	٪۹۰	٪۸۵



شکل ۶- تفاوت عملکرد روش پیشنهادی بر مبنای عامل هوش جمعی

جدول (۵)، خطای میانگین مجذور و انحراف معیار را برای طرح‌های پیشنهادی و روش‌های LSTM و شبکه‌های عصبی بازگشتی استاندارد نشان می‌دهد که نشان از برتری روش پیشنهادی دارد. اگرچه به دلیل واردکردن دانش بیشتر در طرح دوم پیشنهادی این طرح خطای کمتری دارد.

جدول ۶ جزئیات نتایج را با روش‌های معرفی‌شده دیگری که صرفاً با نتایج قضایی کارکرده‌اند مقایسه شده‌اند. مرجع [۲۴] از مفهوم یادگیری عمیق استفاده کرده است و مرجع [۶] نیز از روش شبکه‌های عصبی معمولی استفاده نموده است. همان‌طور که دیده می‌شود، میزان درستی پیش‌بینی‌شده نشان از برتری خوب روش پیشنهادی در هر دو طرح می‌دهد. طرح دوم کمی (حدود ۵ درصد) بهتر از طرح اول رفتار می‌کند؛ اما به یاد داشته باشیم که همین میزان کم می‌تواند در زندگی یک فرد مؤثر باشد.

مزیت اصلی سیستم پیشنهادی فراهم کردن یک ساختار کلی است که در آن بیشتر عوامل مؤثر در پرونده را در نظر می‌گیرد تا به یک تصمیم صحیح با اعتبار کافی برسد. رأی توصیه‌شده، رأیی است که حداکثر جوانب مسئله در آن سنجیده و در نظر گرفته‌شده باشد که ممکن است همه آن‌ها بار اول موردتوجه قاضی قرار نگیرند. بنابراین، می‌تواند همچون یک مشاور به قاضی کمک کند.

جدول ۵- مقایسه میزان خطای سیستم تصمیم‌یار پیشنهادی و روش‌های دیگر

	روش پیشنهادی طرح دوم	روش پیشنهادی طرح اول	LSTM	شبکه‌های عصبی بازگشتی استاندارد
خطا (MSE)	٪۹	٪۱۱	٪۲۱	٪۲۵
STD	۱/۸	۲/۳	۴/۷	۶/۲

جدول ۶- مقایسه عددی میزان درستی روش‌ها

	روش پیشنهادی طرح دوم	**روش پیشنهادی طرح اول**	**سیستم فازی**	شبکه عصبی[۶]	یادگیری عمیق [۲۴]
میزان درستی	٪۹۰/۰۰	٪۸۵/۰۰	٪۷۴/۶۰	٪۴۷/۶۱	٪۷۶/۸۷

۵- نتیجه‌گیری

موضوع دادرسی پرونده‌های قضایی موضوع مهم و نیاز روز است. مسئله طولانی شدن رأی قضایی باعث نارضایتی و ناامیدی از احقاق حق و عدم اجرای کامل عدالت می‌شود. اهمیت این موضوع‌ها از یک‌سو و کاستی‌های پژوهش‌های پیشین از سوی دیگر انگیزه انتخاب این مورد مطالعاتی در این مقاله بوده است. در این راستا، یک سیستم عدالت‌یار هوشمند بر اساس biLSTM پیشنهادشده است. به‌طور مشخص‌تر، با اضافه کردن دولایه بر اساس بهره‌گیری از داده‌های آماری و مفهوم هوش جمعی خبرگان و استفاده از آن برای تمرکز بر مؤلفه‌های مؤثرتر سیستم عدالت‌یار هوشمند پیشنهادشده است. هدف بررسی این مطلب بوده که چگونه با در نظر گرفتن میزان دانش موجود خبرگان که مربوط به عوامل مؤثر در تصمیم‌گیری پرونده‌ها است و با توجه به ساختار دانش خبرگان می‌توان تصمیم‌های مطمئن‌تری داشت؛ درحالی‌که با داده‌های آماری به ارزیابی عوامل مؤثر در این روند پرداخته می‌شود. این امر، تبیین یک فرضیه کلی برای تصمیم‌گیری و قضاوت درباره یک پرونده است. بنابراین، سیستم عدالت‌یار هوشمند پیشنهادی ارائه‌شده می‌تواند یک مشاور و همیار قاضی باشد که به قاضی درگرفتن تصمیم درست کمک می‌کند و با برطرف شدن ابهامات از احقاق حق، نیاز به دادگاه‌های تجدیدنظر را کم می‌کند. درنتیجه، سیستم پیشنهادی می‌تواند راه‌کار مؤثری برای حل مسئله‌ی طولانی شدن رسیدگی‌های قضائی و به عبارتی اطاله‌ی دادرسی باشد.

۶- مراجع

[1] H. A. Dehabadi "The Necessity of Philosophy of Criminology," *fegh and hoghogh* (in Persian), vol. 11, pp. 5-6, 2006.

[2] J. Drobak, N. and C. North, D., "Understanding judicial decision-Making: The importance of constraints on non-rational deliberations," *J. Law Policy*, vol. 134, pp. 26-131, 2008.

[3] C. Guthrie, J. J. Rachlinski, and A. J. Wistrich, "Blinking on the bench: How judges decide cases," *Cornell Law Review*, Forthcoming; Vanderbilt Public Law Research Paper No. 07-25; Vanderbilt Law and Economics Research Paper No. 07-32, 2007.

[4] G. Gigerenzer, Gut feelings: *The intelligence of the unconscious*: Viking Adult, 2007.

[5] M. M. Hall , D. Calabro , T. Sourdin , A. Stranieri , and J. Zeleznikow, "Supporting Discretionary Decision-Making with Information Technology: A Case Study in the Criminal Sentencing Jurisdiction," *University of Ottawa Law and Technology Journal* vol. 2, pp. 1-36, 2005.

[6] V. Aleven "Using background knowledge in case-based legal reasoning: A computational model and an intelligent learning environment," *Artif Intell*, vol. 150, pp. 183-237, 2003.
[7] C. Tata, "The Application of Judicial Intelligence and 'Rules' to Systems Supporting Discretionary Judicial Decision-Making," *in Judicial Applications of Artificial Intelligence*, G. Sartor and K. Branting, Eds., ed Dordrecht: Springer Netherlands, pp. 99-126, 1998.
[8] A. Valente, J. Breuker, and B. O. B. Brouwer, "Legal modeling and automated reasoning with ON-LINE," *International Journal of Human-Computer Studies*, vol. 51, pp. 1079-1125, 1999/12/01/ 1999.
[9] A. Wyner, "An ontology in OWL for legal case-based reasoning," *Artificial Intelligence and Law*, vol. 16, pp. 361-387, 2008/12/01 2008.
[10] J. Keppens, "Argument diagram extraction from evidential Bayesian networks," *Artificial Intelligence and Law*, vol. 20, pp. 109-143, 2012.
[11] C. Li, Y. Hu, and Z. Zhong, "An event ontology construction approach to web crime mining," *in 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 2441-2445, 2010.
[12] T. Abraham and O. d. Vel, "Investigative profiling with computer forensic log data and association rules," *in Proceedings of IEEE International Conference on Data Mining*, pp. 11-18, 2002.
[13] S. Thammaboosadee and U. Silparcha, "A framework for criminal judicial reasoning system using data mining techniques," *in 2nd IEEE International Conference on Digital Ecosystems and Technologies*, pp. 518-523, 2008.
[14] S. C. Naves, "Analytic hierarchy process integrated hybrid agent system for intelligent legal assistance," *in 2011 2nd International Conference on Intelligent Agent & Multi-Agent Systems*, pp. 41-45, 2011.
[15] P. Wang and X. Zhu, "The Reliability Assessment of Legal Reasoning with Evidence Theory," *in Second International Workshop on Knowledge Discovery and Data Mining*, pp. 582-585, 2009.
[16] E. H. Haeusler, V. d. Paiva, and A. Rademaker, "Intuitionistic Description Logic and Legal Reasoning," *in 22nd International Workshop on Database and Expert Systems Applications*, pp. 345-349, 2011.
[17] F. Sabahi and M.-R. Akbarzadeh-T, "Introducing validity in fuzzy probability for judicial decision-making," *International Journal of Approximate Reasoning*, vol. 55, pp. 1383-1403, 2014.
[18] F. Sabahi, *On the Theory of Imprecise Logic and Reasoning based on Extended Fuzzy Logic: Application to Judicial and Medical Decision Making*, Partial Fulfillment of the Requirements of the Degree Doctor of Philosophy (Control Engineering), Faculty of Engineering, Ferdowsi University of Mashhad, 2013.
[19] M. M. J. Theresa and V. J. Raj, "Analogy making in criminal law with neural network," *in International Conference on Emerging Trends in Electrical and Computer Technology*, pp. 772-775, 2011.
[20] K. Zhang, X. Geng, and X. Yan, "Prediction of 3-D Ocean Temperature by Multilayer Convolutional LSTM," *IEEE Geoscience and Remote Sensing Letters*, pp. 1-5, 2020.
[21] J. Han, H. Liu, M. Wang, Z. Li, and Y. Zhang, "ERA-LSTM: An Efficient ReRAM-Based Architecture for Long Short-Term Memory," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, pp. 1328-1342, 2020.
[22] S. D. Kumar and D. Subha, "Prediction of Depression from EEG Signal Using Long Short Term Memory(LSTM)," *in 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1248-1253, 2019.
[23] S. Ye, J. Jiang, J. Li, Y. Liu, Z. Zhou, and C. Liu, "Fault Diagnosis and Tolerance Control of Five-Level Nested NPP Converter Using Wavelet Packet and LSTM," *IEEE Transactions on Power Electronics*, vol. 35, pp. 1907-1921, 2020.
[24] B. Chen, Y. Li, S. Zhang, H. Lian, and T. He, "A Deep Learning Method for Judicial Decision Support," *in IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, pp. 145-149, 2019.
[25] J. L. Kane "Judging credibility," *Litigation Mag.*, Am. Bar Assoc, vol. 33, pp. 1-8, 2007.
[26] L. A. Zadeh, "Probability theory and fuzzy logic are complementary rather than competitive," *Technometrics*, vol. 37, 1995.
[27] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, pp. 157-166, 1994.
[28] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural computation*, pp. 1735-1780, 1997.
[29] A. Graves "Generating Sequences With Recurrent Neural Networks," *arXiv:1308.0850v5*, 2014.
[30] A. Graves, "Generating Sequences With Recurrent Neural Networks," *arXiv:1308.0850v5*, 2014.
[31] Y. Liu, C. Sun, L. Lin, and X. Wang, "Learning Natural Language Inference using Bidirectional LSTM model and Inner-Attention," preprint: *arxiv.org/abs/1605.09090v1*, pp. 1-5, 2016.
[32] L. Cronbach, "Coefficient alpha and the internal structure of tests," *Psychometrika*, vol. 16, pp. 297-334, 1951.
[33] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to Sequence Learning with Neural Networks," *arXiv preprint arXiv:1511.08308*, 2014.
[34] W. Zaremba and I. Sutskeve, "Learning to Execute," *arXiv preprint arXiv:1410.4615*, 2014.
[35] J. Pennington, R. Socher, and M. C. D, "Glove: Global vectors for word representation," *EMNLP*, vol. 14, pp. 1532-1543, 2014.
[36] L. Mou, M. Rui, G. Li, Y. Xu, L. Zhang, R. Yan, et al., "Recognizing entailment and contradiction by tree-based convolution," *arXiv preprint arXiv:1512.08422.*, 2015.

فرناز صباحی مدرک دکترای خود را در رشته مهندسی برق گرایش کنترل از دانشگاه فردوسی مشهد، دریافت کرده‌اند. وی در حال حاضر استادیار و مدیر آزمایشگاه محاسبات نرم در دانشکده فنی و مهندسی، دانشگاه ارومیه می‌باشند. زمینه تحقیقاتی ایشان شامل زمینه‌های منطق و کنترل فازی، سیستم‌های تصمیم‌گیری، رایانش نرم، رباتیک، کنترل هوشمند و یادگیری ماشین می‌باشد.

آدرس پست الکترونیکی ایشان عبارت است از:

f.sabahi@urmia.ac.ir

[1] Inner-Attention
[2] Possibility
[3] Cell State
[4] Gates
[5] Sigmoid
[6] Average pooling
[7] Softmax
[8] Back propagation Trough Time
[9] Cross-Entropy Loss

The CSI Journal on Computing Science and Information Technology

The CSI Journal on
Computing Science and
Information Technology
Vol. 17, No.2, 2020

Intelligent Justice Investigator using Recurrent Neural Network

Farnaz Sabahi

Department of Electrical Engineering, Faculty of Engineering, Urmia University, Urmia, Iran

Abstract

This paper offers a new architecture for practical problems by expanding the use of bilateral short-long memory networks (biLSTM). BiLSTM can entirely reflect the past and future processes of the features. The proposed system has been applied to a judicial case study. The proposed justice investigator and consultant system, for more effective judging, has used two encoders after biLSTM and inner-attention modified by the knowledge of experts. In this architecture, by viewing the folder's features, the type of them is used in the correction of weight. The approach is proposed in two different plans. In each plan, first, biLSTM is applied to the features of the folder in two parts of acquitted and commitment. Accuracy of performance is characterized by the focus on more effective features. The design of this architecture is based on the sharing of weights at the time of training by encoders. In the first plan, we apply the attention factor and then collective intelligence is concerned in the next layers. In the second plan, the collective intelligence is applied in the form of a fuzzy membership function. The results of the proposed consultant system have been compared with alternative approaches in which the superiority of the proposed system is shown. The proposed architecture designs a suitable pattern and employs factors by recognizing their influence in getting the right decision in a shorter time. Therefore, it can be very helpful due to reducing the costs of the appeals and trial and increasing the sense of the trust of community to the judicial system.

Keywords: Decision-making, Judicial Cases, Bilateral LSTM.