

بررسی عملکرد انواع شبکه عصبی بازگشتی در پیش‌بینی داده‌های سری زمانی

در حمل و نقل؛ نوع داده: مسیر حرکت عابرپیاده در پیاده‌رو

مقاله علمی - پژوهشی

*علی ادیسی (نویسنده مسئول)، دانشیار، دانشکده مهندسی عمران، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران

محمد زاهدی، دانشجوی کارشناسی ارشد، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران

*پست الکترونیکی نویسنده مسئول: edrisi@kntu.ac.ir

دریافت: ۱۴۰۴/۰۴/۰۹ - پذیرش: ۱۴۰۴/۰۸/۲۰

صفحه ۱۲۴-۱۰۷

چکیده

در دنیای امروز، هوش مصنوعی به عنوان یک عامل قدرتمند غیرقابل انکار در تمام امور انسانی ورود پیدا کرده است. این هوش با استفاده از الگوریتم‌های شبکه عصبی و کلان داده‌ها آموزش دیده و پیش‌بینی می‌کند. در گذر زمان انواع متفاوتی از شبکه‌های عصبی برای کاربردهای متفاوت معرفی و توسعه داده شده که یکی از اینها، شبکه عصبی بازگشتی (RNN) است. مدل RNN بخاطر معماری و ساختار آن، عملکرد قابل قبولی بر روی داده‌های سری زمانی، داده‌هایی که در آن ویژگی‌های مختلف از یک پدیده در گام‌های زمانی ثابت برداشت شده- دارد. در حمل و نقل موارد زیادی از داده‌های سری زمانی وجود دارد مانند: حجم عبوری ترافیک از یک نقطه خاص در بازه‌های زمانی ثابت، تعداد و جنسیت و ... مسافرین مترو در ساعات مختلف شبانه‌روز، ویژگی‌های حرکت مثل موقعیت و سرعت و شتاب یک عامل ترافیکی مثل عابر پیاده رو هر لحظه در پیاده رو. مورد آخر که سوابق و تاریخچه ویژگی‌های حرکتی عابر است، اساس کار این پژوهش است. در این مقاله با استفاده از ۳ زیرمجموعه شبکه عصبی بازگشتی یعنی مدل‌های *Vanilla LSTM*، *Stacked LSTM* و *GRU* به پیش‌بینی مسیر حرکت عابر می‌پردازیم. ۲ هدف اصلی این پژوهش اولاً تخمین موقعیت آینده یک عابر جهت شناسایی و رفع شرایط خطرآفرین در تعامل عابر و سیستم خودران (مثل ربات کالارسان) در پیاده رو و ثانیاً بررسی عملکرد این ۳ مدل در داده‌های سری زمانی است. نتایج نشان داد که در افق زمانی پیش‌بینی کوتاه مدت، مدل *GRU* عملکرد بهتری نسبت به دیگر مدل‌ها دارد. اما با افزایش افق زمانی پیش‌بینی و افزایش پیچیدگی‌های داده‌ها، مدل *Stacked LSTM* عملکرد بهتری نسبت به سایرین دارد.

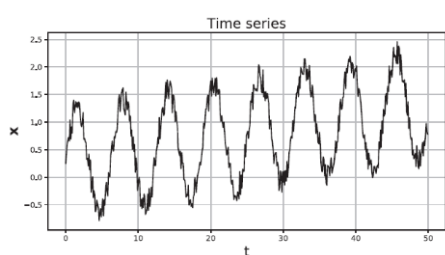
واژه‌های کلیدی: پیش‌بینی مسیر حرکت، داده‌های سری زمانی، شبکه عصبی بازگشتی، LSTM، GRU

۱- مقدمه

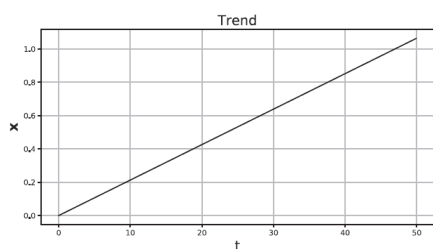
سری‌های زمانی، استفاده از داده‌های جمع‌آوری شده از مشاهدات گذشته برای ایجاد یک مدل دقیق است که الگوی موجود در سری‌ها را منعکس کند و پیش‌بینی و طبقه‌بندی رویدادهای آینده را امکان‌پذیر سازد. اهمیت این موضوع، تحلیل سری‌های زمانی را به بخش اساسی مدل‌سازی داده‌ها در طیف وسیعی از صنایع تبدیل کرده است (Weerakody et al, 2021). یک سری زمانی چند متغیره، دنباله‌ای از مشاهدات است که در طول زمانی

در مهندسی حمل و نقل، روش‌های گوناگونی برای جمع‌آوری و ارائه اطلاعات معرفی و توسعه داده شده است. یکی از این روشها، داده‌های سری زمانی است. سری‌های زمانی به مجموعه‌ای از مشاهده ویژگی‌های یک پدیده که به ترتیب زمانی مرتب شده‌اند، اشاره دارند. تجزیه، تحلیل و پیش‌بینی سری‌های زمانی می‌تواند یک پشتیبان قوی برای تصمیم‌گیری در اختیار ما قرار دهد (Hu et al, 2020). هدف از تحلیل

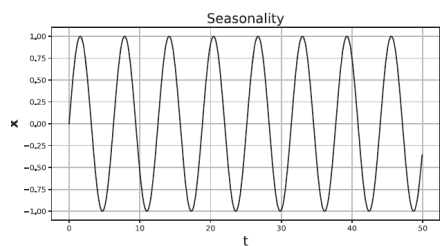
سری منظم معمولاً دارای ۲ مشخصه اصلی است (شکل ۱)، الف) روند که حرکت کلی ای است که سری زمانی در طول دوره مشاهده از خود نشان می‌دهد و در برخی متون، این مؤلفه به عنوان تغییرات بلندمدت نیز شناخته می‌شود و ب) نوسانات فصلی (دوره‌ای) که این مؤلفه تغییراتی را که در فواصل منظم خاص رخ می‌دهند، شناسایی می‌کند و اطلاعات مفیدی را در مورد دوره‌های زمانی با الگوهای مشابه ارائه دهد. از سوی دیگر با رشد سیستم‌های چند حسگری و همچنین استفاده مداوم از داده‌های بدون ساختار، سری‌های زمانی نامنظم به طور فزاینده‌ای رایج شده‌اند (Tores et al 2020, Weerakody et al, 2021).



الف



ب



پ

شکل ۱. الف) تغییرات متغیر x بر حسب زمان که از نوع سری زمانی منظم است. ب) روند کلی تغییرات (پ) نوسانات فصلی (دوره‌ای)

برای تحلیل سری‌های زمانی الگوریتم‌های مختلفی معرفی شده است. روش‌های اولیه مانند مدل‌های AR^2 ، MA^3 و $ARMA^4$ برای داده‌های منظم مناسب بودند، اما با الگوهای نامنظم مطابقت

متوالی، با گام‌های مساوی، ثبت شده‌اند. برخلاف سری تک متغیره (که یک متغیر را در طول زمان دنبال می‌کند)، سری چند متغیره چندین متغیر مرتبط را به طور همزمان ثبت می‌کند و این امکان را می‌دهد وابستگی‌های زمانی و تعاملات بین سری‌ها را مدل‌سازی کنیم.

مشاهده و برداشت اطلاعات ویژگی‌های مربوط به مسیر حرکت یک عابر پیاده را می‌توان به عنوان داده سری زمانی چندمتغیره در نظر گرفت. در این پدیده، موقعیت مکانی، سرعت و شتاب، ویژگی‌های مختلف هستند که در گام‌های زمانی ثابت، برداشت می‌شوند. با ورود سیستم‌های خودران به محیط‌های انسانی (مانند ورود ربات کالارسان به پیاده رو)، اهمیت درک و پیش‌بینی رفتار انسان توسط این سیستم‌ها روز به روز بیشتر می‌شود. بطور خاص، پیش‌بینی موقعیت آتی کاربران و برنامه‌ریزی و واکنش برای چنین پیش‌بینی‌هایی جزء وظایف کلیدی برای وسایل خودران است. پیش‌بینی مسیر حرکت انسان به دلیل پیچیدگی رفتار و محرک‌های مختلف درونی و بیرونی، کاری سخت است. پارامترهایی مانند هدف، محیط، روابط نامرئی اجتماعی، قوانین و غیره در این پیش‌بینی نقش دارند (Rudenko et al. 2020).

در این پژوهش با به‌کارگیری روش‌های پیشرفته شبکه عصبی بازگشتی^۱ (RNN)، به پیش‌بینی موقعیت آینده عابران پرداخته شده و دقت این پیش‌بینی‌ها از طریق معیارهای معتبر سنجیده می‌شود. همچنین، الگوریتم‌های معرفی شده در تحلیل و پیش‌بینی داده‌های سری زمانی با یکدیگر مقایسه شده و نقاط قوت و ضعف هر کدام در شرایط مختلف مورد ارزیابی قرار می‌گیرد.

۲- پیشینه تحقیق

یکی از حوزه‌های اصلی یادگیری ماشین، پیش‌بینی سری‌های زمانی است، زیرا بسیاری از مسائل مختلف مانند تجارت، مدیریت زنجیره تأمین و تولید شامل یک مؤلفه زمان هستند. مجموعه داده‌های سری زمانی با سایر مجموعه داده‌ها متفاوت است، زیرا یک عنصر زمان را اضافه می‌کند. این عنصر اضافه شده هم یک محدودیت و هم یک منبع اطلاعات اضافی را ارائه می‌دهد (Yamak et al, 2020). یک سری زمانی استاندارد دنباله ای از T نقطه داده است که به ترتیب زمانی مرتب شده‌اند و به صورت $X = \{x_1, x_2, x_3, \dots, x_T\} \in \mathbb{R}^{T \times D}$ که در آن $x_t \in \mathbb{R}^D$ نقطه داده در گام زمانی t و D بُعد هر نقطه داده است.

اگر $D = 1$ آنگاه سری زمانی تک متغیره است و اگر $D > 1$ آنگاه سری زمانی چندمتغیره است (liang et al, 2024). داده سری زمانی را از نظر نوع تغییرات مقادیر متغیرهای آن می‌توان به ۲ فرم کلی منظم و نامنظم تقسیم‌بندی کرد. یک داده

سالخوردگان و افراد ناتوان، مسیریابی بهینه و کاهش مصرف سوخت و امکان استراحت برای همه‌ی سرنشینان نام برد (Hansson et al, 2021). اما ورود خودران‌ها با تهدیدهایی نیز همراه است. یک تهدید، عملکرد نامناسب سیستم‌های کامپیوتری یا سنسورهای آن در لحظات حساس است که می‌تواند باعث یک فاجعه شود. نگرانی دیگر نحوه تعامل بین خودران‌ها و خودروهایی فعلی است. اینکه خودران‌ها در هر لحظه بتوانند ارتباط و موقعیت سایرین را درک کنند، یک چالش مهم است. نگرانی مهم دیگر ایمنی رایانه‌ای و جرائم مربوطه است. امکان هک شدن سیستم‌های خودران و کنترل کردن آنها توسط افراد تبهکار می‌تواند باعث ایجاد بحث‌های جدی در حوزه امنیتی باشد (Gruel and Standford 2015, Liu et al, 2018, ji, 2018).

جدول ۱. الف) مشخصات ۳ مدل استفاده شده در پژوهش آمالو و همکاران (۲۰۲۲) ب) نتایج پژوهش برای ۳ شاخص مختلف

الف)			
	RNN	LSTM	GRU
Hidden layers	۳	۱	۴
Number of units in a hidden layer	۸۰	۱۰۰	۵۰
Epochs	۲۰	۲۰	۲۰
ب)			
	RNN	LSTM	GRU
RMSE	۰/۰۵۱	۰/۰۳۹	۰/۰۳۴
MAE	۰/۰۴۱	۰/۰۲۱	۰/۰۲۲
R2-Score	۰/۲۱	۰/۵۳	۰/۶۴

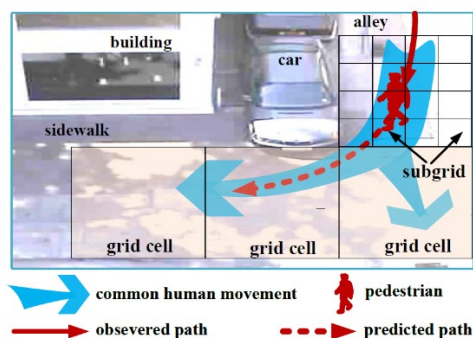
پیشبینی مسیر حرکت عابران پیاده که به عنوان یک پدیده سری زمانی چندمتغیره با مدلسازی رگرسیونی شناخته می‌شود، توجه بسیاری از پژوهشگران را در سال‌های به خود جلب کرده است. در این پیش بینی به دلیل پیچیدگی رفتار انسان و امکان تغییر تصمیم (مسیر) در لحظه، پژوهشگران را به توسعه‌ی مدل‌هایی با دقت بیشتر و استفاده از رویکردهای گوناگون سوق داده است. این رویکردها بر عواملی مانند تعامل میان عابران پیاده، اهداف حرکتی و مقاصد آن‌ها تمرکز دارند تا بتوانند الگوهای حرکتی را با خطای کمتری پیش‌بینی کنند. هدف اصلی، طراحی مدل‌هایی است که بتوانند رفتارهای واقعی افراد را به درستی بازتاب دهند

نداشتند. روش $ARIMA^5$ برخی از این محدودیت‌ها را برطرف کرد، با این حال اکثر پدیده‌های دنیای واقعی غیرخطی هستند که همین موضوع باعث گسترش استفاده از مدل‌های یادگیری ماشین شده‌اند. تکنیک‌هایی مانند ماشین‌های بردار پشتیبان^۶ (SVM) و شبکه‌های عصبی بازگشتی (RNN) - به دلیل توانایی‌شان در یادگیری الگوها به طور مستقیم از داده‌ها، فراگیرتر شده‌اند. در میان این‌ها، معماری‌های پیشرفته‌ای مانند مدل حافظه بلند-کوتاه مدت^۷ (LSTM) و واحد دروازه بازگشتی^۸ (GRU) به دلیل قابلیت‌های یادگیری توالی و داشتن حافظه برجسته هستند که منجر به دقت پیش‌بینی بالا در زمینه‌هایی مانند پیش‌بینی انرژی، مدلسازی ترافیک، پیش‌بینی محیطی و غیره می‌شوند (Hu et al, 2020).

همانطور که پیشتر گفته شد، سری‌های زمانی مختلفی در حمل و نقل وجود دارد. پژوهشی درباره پیش بینی جریان ترافیک در دو حالت تک متغیره (جریان ترافیک: میانگین تعداد وسایل نقلیه شناسایی شده در هر دقیقه در یک ایستگاه خاص) و چندمتغیره (۱. جریان ترافیک، ۲. سرعت: میانگین سرعت وسایل نقلیه عبوری از ایستگاه و ۳. اشغال: درصد زمانی که حسگر با وسیله نقلیه مشغول است) نشان داد که یک مدل LSTM چند متغیره به دقت پیش‌بینی بالاتری (شاخص‌های MAE و RMSE پایین‌تر) نسبت به مدل‌های تک متغیره دست می‌یابد (Kumar and Hariharan, 2020). در پژوهش دیگری مدلی براساس مکانیسم توجه^۹ مبتنی بر ویژگی‌ها (و نه صرفاً زمان) برای پیش‌بینی سری‌های زمانی چند متغیره، به منظور ثبت وابستگی بین متغیرها در طول زمان استفاده شد. نتایج عملکرد قابل قبولی را هم بر روی داده‌های آزمایشگاهی و هم بر روی داده‌های دنیای واقعی نشان داد (Shih et al, 2019). در پژوهشی دیگر با بررسی مصرف انرژی برق به عنوان یک داده‌ی سری زمانی چندمتغیره، مطابق جدول ۱ مشاهده گردید که در شاخص‌های RMSE و R2-Score مدل GRU به مراتب از مدل‌های LSTM و simple RNN عملکرد بهتری دارد (Amalou et al, 2022).

در سال‌های اخیر با فراگیر شدن استفاده از سیستم‌های خودران در محیط‌های انسانی، بررسی موقعیت آینده و پیش‌بینی مسیر حرکت عابرین پیاده از اهمیت زیادی برخوردار شده است. از نتایج مطلوب ورود خودران‌ها به شبکه ترافیک می‌توان به حذف رانندگی پرخطر، افزایش دسترسی به خودرو برای

و حرکت عابرین (Pedestrian-LSTM) بود که با لحاظ کردن اطلاعات اجسام ساکن در محیط آموزش دید. در این کار یک ساختار دو شبکه‌ای دو سطحی (سلول‌های شبکه و زیرشبکه) روی هم قرار داده شد (Huynh and Alagband, 2019).



شکل ۳. Scene-LSTM جابه‌جایی عابر را در ساختار شبکه‌ای دو سطحی یاد می‌گیرد. جابه‌جایی رایج عابرین در ترکیب با حرکت فردی (Pedestrian-LSTM) برای پیش‌بینی مکان‌های آینده عابران استفاده می‌شود. (Huynh and Alagband, 2019)

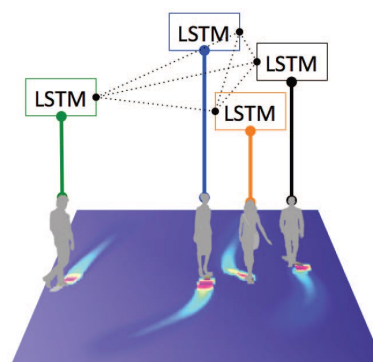
مطالعات بالا که بخش کوچکی از پژوهش‌های حوزه داده سری زمانی و پیش‌بینی حرکت بود، نشان از گستردگی موضوع و تنوع روشها برای تحلیل و تخمین موقعیت کاربر دارد. در میان این روشها، شبکه‌های عصبی به ویژه شبکه عصبی بازگشتی و زیرمجموعه‌های آن یعنی مدل LSTM و GRU به دلیل داشتن حافظه بلندمدت در ساختار خود، توانایی بالایی در یادگیری الگوها و یادآوری اطلاعات گذشته دارند. برای پیش‌بینی مسیر حرکت در این پژوهش، ۳ مدل Vanilla LSTM و Stacked LSTM (S-LSTM) و GRU استفاده شده و با تغییر شرایط و پارامترهای هر مدل، به بررسی عملکرد هر کدام در پیش‌بینی مسیر حرکت می‌پردازیم.

۳- روش تحقیق

در این قسمت و در ۴ بخش به معرفی روش انجام پژوهش می‌پردازیم. بخش ۳-۱ به معرفی مختصری از شبکه‌های عصبی اختصاص دارد. بخش ۳-۲ به معرفی شبکه LSTM، GRU و S-LSTM، بخش ۳-۳ به معرفی جزئیات مدل، بخش ۳-۴ به معرفی دیتاست، نرم افزار و سخت افزار استفاده شده و بخش

و پیش‌بینی‌های دقیق‌تری ارائه دهند. به‌طور کلی، فرآیند پیش‌بینی در این رویکردها شامل دو مرحله‌ی اساسی است: ۱. استخراج ویژگی‌ها: شناسایی مجموعه‌ای از عوامل مؤثر بر حرکت افراد. ۲. برازش مدل: تطبیق مدل با داده‌های رفتاری برای دستیابی به نتایج دقیق‌تر (Bighashdel and Dubbelman, 2019).

در سال‌های اخیر، شبکه‌های عصبی بازگشتی، به‌ویژه مدل LSTM، به یکی از روش‌های پرکاربرد در یادگیری دنباله‌ها و پیش‌بینی حرکت تبدیل شده‌اند. آلهی و همکارانش از نخستین پژوهشگرانی بودند که مدل Social-LSTM را برای پیش‌بینی مسیرهای هم‌زمان در محیط‌های پیوسته پیشنهاد کردند. در این روش، حرکت هر فرد با استفاده از یک واحد LSTM جداگانه مدل‌سازی می‌شود. از آن‌جا که رفتار عابران تحت تأثیر افراد اطرافشان قرار دارد، این واحدهای LSTM در قالب یک سامانه تعامل اجتماعی به هم متصل می‌شوند و از طریق حالت پنهان خود، داده‌ها را با عابران نزدیک به اشتراک می‌گذارند (Alahi et al, 2016).



شکل ۲. هدف مقاله (Alahi et al, 2016) پیش‌بینی مسیر حرکت در مکان‌های شلوغ است که حرکت هر فرد تحت تأثیر اطرافیان است.

در این پژوهش یک مدل جدید به نام Social LSTM معرفی شد که با در نظر گرفتن قوانین منطقی و قراردادهای اجتماعی که انسان‌ها هنگام پیمایش در محیط‌های مشترک از آنها استفاده می‌کنند، مسیر همه افراد در یک صحنه را پیش‌بینی می‌کند.

بارتولی و همکاران (۲۰۲۰) با توسعه‌ی مدل Social-LSTM، یک ساختار کلی‌تر مبتنی بر ادغام زمینه‌محور^{۱۲} ارائه دادند که قادر است تأثیر عناصر محیطی بر مسیر حرکت افراد را تخمین بزند (Bartoli et al, 2018). هیون و آلاکباند یک مدل LSTM براساس شبکه‌بندی محیط حرکت عابرین معرفی کردند. مدل آنها دربرگیرنده اطلاعات محیط (Scene-LSTM)

بهینه‌سازی مانند گرادیان کاهشی تصادفی^{۱۱} (SGD) به صورت پیاپی وزن‌های مدل را بروزرسانی می‌کنند. محور این فرآیند، الگوریتم پس‌انتشار^{۱۲} است که با انتشار خطا به عقب از طریق لایه‌های شبکه، گرادیان تابع خطا را نسبت به هر وزن محاسبه می‌کند. این گرادیان‌ها جهت و بزرگی تغییر مورد نیاز برای هر پارامتر را نشان می‌دهند و به بهینه‌ساز اجازه می‌دهند مدل را اصلاح کرده و دقت پیش‌بینی آن را در طول زمان بهبود بخشد (Zhao et al, 2022).

شبکه‌های DNN، نمی‌توانستند توالی در سری زمانی را به درستی پردازش کنند. شبکه عصبی بازگشتی (RNN) مدلیست که بازگشت (به نوعی تکرار) را در نظر می‌گیرد. در این مدل، خروجی یک لایه به عنوان ورودی در گام بعدی به مدل داده می‌شود. شبکه‌های بازگشتی قادر به مدل‌سازی داده‌های توالی‌دار به منظور تشخیص و پیش‌بینی هستند. این شبکه‌ها از حالت‌های پنهان با ابعاد بالا و دینامیک غیرخطی ساخته شده‌اند. حالت‌های پنهان (h در شکل ۵) به عنوان حافظه شبکه کار می‌کنند و وضعیت آنها در هر زمان مشروط به وضعیت قبلی آن است. این ساختار آنها را قادر می‌سازد تا داده‌های قبلی را برای دوره‌های زمانی طولانی ذخیره و پردازش کنند. این شبکه‌ها می‌توانند یک توالی ورودی را به یک توالی خروجی در گام زمانی فعلی نگاشت کنند و توالی جدید برای گام زمانی بعدی پیش‌بینی کنند (Salehnejad et al, 2018).

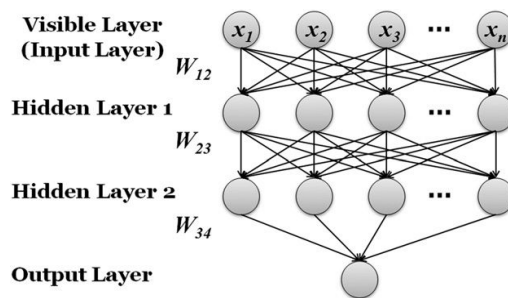
۳-۲-مدل LSTM، S-LSTM و GRU

مشکل اصلی RNN این است که با افزایش فاصله زمانی، مستعد انفجار گرادیان^{۱۳} یا محو گرادیان^{۱۴} است. به منظور حل این مشکل، مکانیسم آستانه برای کنترل انباشت اطلاعات معرفی شد که مدل LSTM اصلی‌ترین مدل برای این مکانیسم است. ایده LSTM ساده است. همانطور که در شکل ۵ نشان داده شده، لایه پنهان RNN اصلی فقط یک حالت h دارد که با ورودی‌های کوتاه‌مدت سازگار است. بنابراین، یک حالت C به RNN اضافه می‌شود تا حالت بلندمدت را به خاطر بسپارد. این اساس LSTM است (Jiang et al, 2019)

۳-۵ به معرفی روش مدلسازی و شاخص‌های ارزیابی مدل اختصاص دارد.

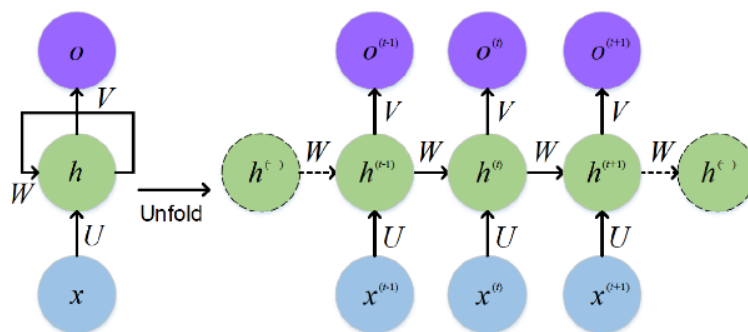
۳-۱- شبکه عصبی عمیق (DNN) و شبکه عصبی بازگشتی (RNN)

شبکه‌های عصبی مصنوعی^{۱۱} (ANN) از لایه‌هایی از گره‌هایی متصل به نام نورون‌های مصنوعی ساخته شده‌اند. یک «شبکه کم‌عمق» به یک ANN با یک لایه ورودی، یک لایه خروجی و حداکثر یک لایه پنهان بدون اتصال بازگشتی اشاره دارد. با افزودن به تعداد لایه‌ها، عمق و پیچیدگی شبکه افزایش می‌یابد (DNN) و آن را قادر می‌سازد تا به صورت کارآمدتری الگوهای درون کلان داده‌ها را شناسایی کند و یاد بگیرد که اصطلاحاً این یادگیری به «یادگیری عمیق» شناخته می‌شود.

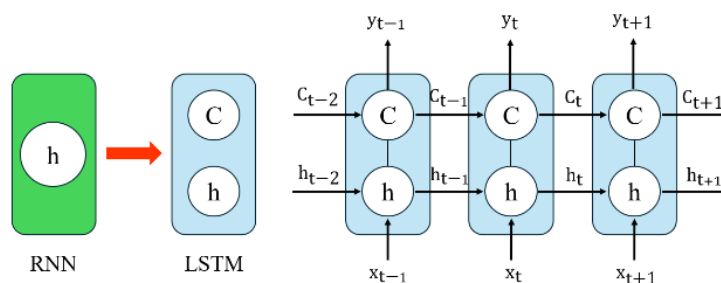


شکل ۴. معماری یک DNN با یک لایه پنهان. بردارهای ورودی در وزن‌های مربوط به خود ضرب شده و به نورون مقصد می‌رسند. سپس مجموع ورودی‌های هر نورون از یک تابع فعال سازی عبور کرده و از نورون خارج می‌شوند. این پروسه در هر لایه تکرار می‌شود تا محاسبات به لایه خروجی برسند (Haykin 2009)

در یادگیری عمیق، آموزش شامل تنظیم پارامترهای یک شبکه عصبی برای به حداقل رساندن تفاوت بین پیش‌بینی‌های آن و مقادیر هدف واقعی است، اختلافی که توسط تابع خطا (یا تابع زیان) اندازه‌گیری می‌شود. برای کاهش این خطا، الگوریتم‌های



شکل ۵. شکل ساده شده از یک سلول بازگشتی (چپ) و جزئیاتی که در آن اتفاق می‌افتد (راست). خروجی هر گام زمانی مجدداً برای پیش بینی گام زمانی بعدی، وارد شبکه می‌شود. U و W و V ماتریس وزن‌های شبکه، به ترتیب برای لایه ورودی به پنهان، پنهان به پنهان و پنهان به خروجی هستند (Feng et al, 2017).



شکل ۶ تفاوت RNN و LSTM در حالت C است.

$$\text{Output gate: } o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (۳)$$

Intermediate Cell State:

$$\tilde{C} = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (۴)$$

Cell State (next memory input):

$$c_t = (i_t * \tilde{C}) + (f_t * c_{t-1}) \quad (۵)$$

$$\text{New h state: } h_t = o_t * \tanh(c_t) \quad (۶)$$

در روابط بالا، σ و \tanh توابع فعالسازی، W و b به ترتیب وزن‌ها و سوگیری^{۱۹} مربوطه برای هر دروازه و x_t داده ورودی در گام زمانی t است.

در مدل LSTM، یادگیری با مقایسه خروجی‌های پیش‌بینی شده با مقادیر واقعی و محاسبه خطا با استفاده از یک تابع زیان انجام می‌شود. سپس این خطا با استفاده از روش پس انتشار در گذر زمان^{۲۰} (BPTT) در جهت عکس در شبکه منتشر می‌شود.

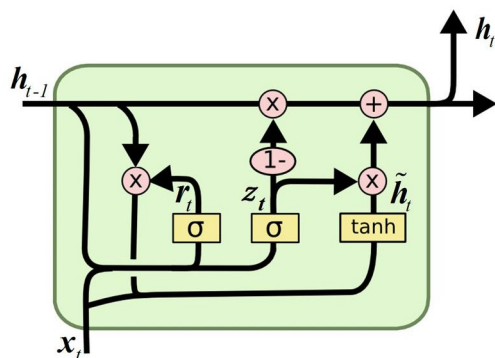
مطابق شکل ۶، در لحظه t سه ورودی برای یک سلول LSTM وجود دارد، مقدار ورودی شبکه در لحظه فعلی x_t مقدار خروجی LSTM در لحظه قبل h_{t-1} و حالت C در لحظه قبل c_{t-1} . خروجی LSTM دارای دو مقدار است، h_t در لحظه فعلی، و حالت C_t در لحظه فعلی. که در آن x و h همگی بردار هستند. سلول LSTM از یک دروازه ورودی^{۱۵}، یک دروازه فراموشی^{۱۶} و یک دروازه خروجی^{۱۷} استفاده می‌کند (شکل ۷). این دروازه‌ها بسته به اولویت داده‌ها شبکه را قادر می‌سازند تا یاد بگیرد چه چیزی را ذخیره کند، چه چیزی را فراموش کند، چه چیزی را به خاطر بسپارد، چه چیزی را مورد توجه قرار دهد و چه چیزی را خروجی دهد. حالت سلول^{۱۸} (C) و حالت پنهان (h) در جمع‌آوری داده‌ها برای پردازش در حالت بعدی استفاده می‌شوند.

بنابراین، از مشکل گرادیان محوشونده جلوگیری می‌شود. دروازه‌ها معادلات زیر را دارند (Yamak et al, 2019).

$$\text{Input gate: } i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (۱)$$

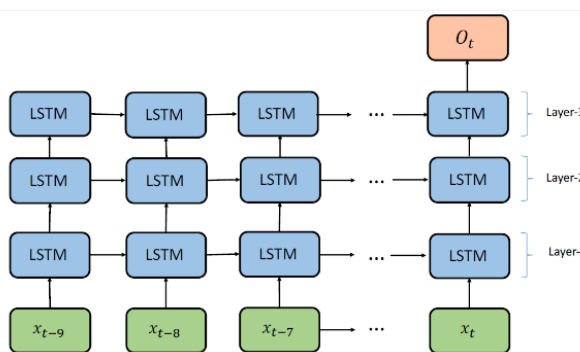
$$\text{Forget gate: } f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (۲)$$

سادگی، مدل‌های GRU به طور گسترده مورد استفاده قرار می‌گیرند و برای ثبت وابستگی‌ها در داده‌های متوالی عملکرد قابل قبولی دارد (Torres et al, 2020).



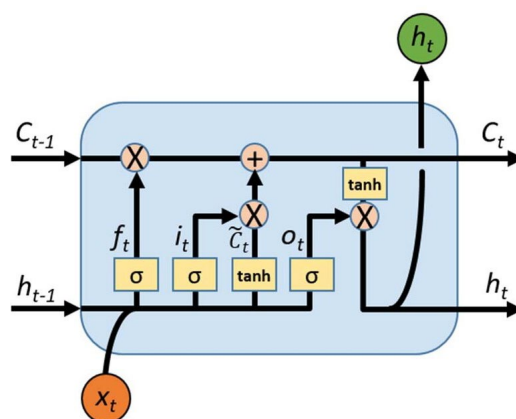
شکل ۸. معماری سلول GRU

ساده‌ترین روش برای افزودن ظرفیت به شبکه LSTM، روی هم چیدن لایه‌های LSTM است. این عمل منجر به تشکیل مدل Stacked LSTM (S-LSTM) شده که ابتدایی‌ترین و ساده‌ترین ساختار شبکه LSTM است. همچنین می‌توان آن را به عنوان یک ساختار چندلایه کاملاً متصل در نظر گرفت. در شبکه S-LSTM، خروجی لایه $(L - 1)$ ام در زمان t ، h_t^{L-1} است. این خروجی به عنوان ورودی، x_t^L لایه L ام در نظر گرفته می‌شود. این اتصالات خروجی-ورودی تنها رابطه بین دو لایه مجاور هم هستند. در بُعد زمان، اتصالات بازگشتی فقط در یک لایه قرار دارند (شکل ۹).



شکل ۹. تصویر ساده ای از S-LSTM. چپ: مدلی با ۳ لایه. راست: باز شده مدل در طول زمان با فرض وجود ۱۰ گام زمانی سابق برای داده‌های ورودی (Saleh et al, ۲۰۱۷)

با بررسی تأثیر هر پارامتر بر خروجی نهایی، مدل وزن‌های خود را برای به حداقل رساندن خطاهای آینده بازتعریف می‌کند. ساختار منحصر به فرد LSTM - به ویژه حالت C و مکانیسم‌های دروازه - به حفظ اطلاعات مهم در توالی‌های طولانی کمک می‌کند و آن را در یادگیری الگوهای زمانی مؤثرتر می‌کند. از طریق تکرار رو به جلو و به‌روزرسانی‌های رو به عقب، مدل به تدریج پیش‌بینی‌های خود را با گذشت زمان بهبود می‌بخشد.



شکل ۷. معماری یک سلول LSTM (Sun et al, 2018). در تصویر علامت * و + به معنای ضرب و جمع نظیر به نظیر بین درایه های ۲ ماتریس است.

مدل GRU نسخه ساده‌شده از LSTM است که برای کاهش پیچیدگی محاسباتی معرفی شد (شکل ۸). در حالی که LSTM از سه دروازه استفاده می‌کند، GRU دارای دو دروازه - دروازه به‌روزرسانی (Z_t) و دروازه تنظیم مجدد (r_t) - است علیرغم

۳-۳- جزئیات مدل

در مدل LSTM مانند سایر مدل‌های شبکه عصبی، پارامترهای زیادی در پیش پردازش داده‌ها، تزریق داده‌ها به مدل، آموزش و بهینه سازی تابع خطا و گرفتن نتایج دخیل هستند که در جدول زیر به معرفی کوتاهی از آنها می‌پردازیم.

Data splitting: جداسازی داده‌ها به آموزش، اعتبارسنجی و ارزیابی

Scaling: فرآیند تبدیل داده‌های عددی به یک محدوده خاص تا یادگیری و محاسبات مدل را آسان‌تر کنند.

Activation function: تابع فعالسازی که وظیفه‌اش ایجاد یک خروجی با شکل و مقدار مشخص است.

Dropout: بطور تصادفی بخشی از نورون‌ها را نادیده می‌گیرد تا از بیش‌برازش جلوگیری کند.

Learning rate: عددی که میزان بروزرسانی وزن‌های مدل را در پاسخ به خطایی که در طول آموزش می‌بیند، کنترل می‌کند.

Batch size: تعداد نمونه‌هایی از داده که قبل از بروزرسانی وزن‌ها همزمان با هم پردازش شده‌اند.

Epoch: تعداد گذر کامل از داده‌های آموزشی.

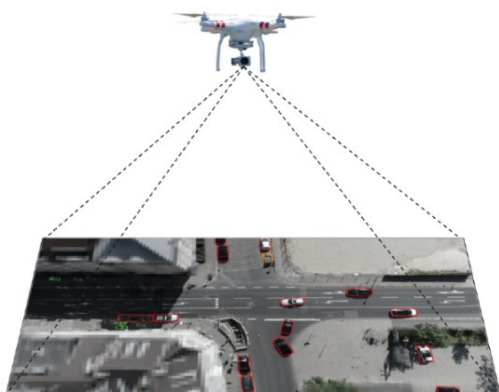
Loss function: تابع زیان که «خطا» را تعریف می‌کند (تفاوت بین پیش‌بینی و واقعیت).

Optimizer: الگوریتمی که وزن‌های مدل را مقداردهی مجدد می‌کند تا تابع زیان را به حداقل رسانده و دقت مدل را بهبود بخشد.

Metrics: معیارهای کمی برای ارزیابی عملکرد یک مدل که نشان می‌دهند پیش‌بینی مدل چقدر دقیق است.

۳-۴- معرفی دیتاست و نرم افزار مورد استفاده

دیتاست inD (inD Dataset) که در این پژوهش مورد استفاده قرار گرفته، از نوع نگاه از بالا به وسیله دوربین نصب شده بر روی پهباد است که ویژگی‌های حرکت شامل موقعیت، سرعت، شتاب و ... در هر گام زمانی (در اینجا هر فریم از فیلم که معادل $\frac{1}{25}$ ثانیه است) برای همه کاربران شبکه ترافیک (عابریاده، دوچرخه و خودروی سواری و سنگین) در زمان حضور آنها در قاب دوربین پهباد ثابت شده (بدون حرکت در آسمان) بر روی ۴ تقاطع مختلف در شهر آخن آلمان است. این دیتاست مجموعه‌ای جامع از اطلاعات کاربران شبکه با رفتار طبیعی است.



شکل ۱۰. دیتاست inD. مسیر حرکت هر کاربر شبکه ترافیک با استفاده از الگوریتم‌های یادگیری عمیق به طور دقیق استخراج می‌شود (Bock et al, 2020).

مجموعه داده حاصل شامل بیش از ۱۳,۵۰۰ کاربر مختلف در ۱۰ ساعت فیلمبرداری در ساعات مختلف شبانه‌روز است.

در inD، برای هر ضبط، پهباد در ارتفاع ۱۰۰ متری قرار گرفت تا به طور کامل تقاطع را پوشش دهد. ویدئوها با DJI Phantom 4 Pro با وضوح 4K (4096*2160 pixel) با حداکثر کیفیت و با سرعت ۲۵ فریم بر ثانیه ضبط شده‌اند. زمان ضبط هر ویدئو بین ۲۰ تا ۲۲ دقیقه است. برای هر کاربر و در هر فریم، ویژگی‌های زیر از حرکتش استخراج شده است.

-موقعیت مکانی نسبت به محور X و Y

-زاویه حرکت (heading) نسبت به محور Y

-ابعاد وسیله (عرض و طول) که برای عابر، صفر است

-سرعت و شتاب نسبت به محور X و Y

-سرعت و شتاب طولی و جانبی

برای تمامی مراحل پژوهش اعم از پیش پردازش داده‌ها، طراحی و آموزش مدل و تفسیر و مصورسازی نتایج از زبان برنامه‌نویسی پایتون و کتابخانه‌های عمومی آن مثل Numpy، Pandas، Matplotlib و کتابخانه‌های ویژه یادگیری ماشین و عمیق مثل Scikit-Learn و TensorFlow استفاده شده است. همچنین از پلتفرم google colab با مشخصات جدول ۲ برای این پژوهش استفاده شد.

جدول ۲. مشخصات سخت افزار استفاده شده

مشخصات	سخت افزار
Intel Xeon (typically 2 virtual cores)	واحد پردازنده مرکزی (CPU)
NVIDIA Tesla T4, P100, or K80	واحد گرافیک (GPU)
12-13 GB	حافظه موقت (RAM)

۳-۵-۳- روش مدل‌سازی و شاخص‌های ارزیابی

روند مدل‌سازی در این پژوهش ۵ گام به ترتیب زیر و شرحی که در ادامه می‌آید انجام شده است: گام اول: پیش پردازش داده‌ها (Preprocessing)، گام دوم: انتخاب نوع مدل، گام سوم: تنظیم پارامترها، گام چهارم: آموزش مدل و گام پنجم: انجام پیش‌بینی و مصورسازی نتایج

۳-۵-۳-۱- گام اول: پیش پردازش داده‌ها

گام اول که انجام پیش پردازش بر روی داده‌های خام اولیه است از مهمترین مراحل در اکثر مدل‌سازی‌ها در یادگیری عمیق می‌باشد.

مقیاس‌دهی: برای مقیاس دهی به داده‌ها که عبارتست از تبدیل مقادیر عددی برای یک ویژگی به محدوده‌ای خاص، معمولاً از ۲ روش کمینه-بیشینه یا استانداردسازی استفاده می‌شود. در این پژوهش از روش اول استفاده شده که داده‌ها را مطابق رابطه (۷) به محدوده ۰ تا ۱ تبدیل می‌کند.

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (7)$$

در رابطه اخیر، x_{min} و x_{max} به ترتیب مقادیر کمینه و بیشینه از ویژگی x (فرضاً سرعت عابر) و x_i مقدار ویژگی در لحظه $t = i$ و x'_i مقدار مقیاس شده x_i است.

انتخاب ویژگی‌های ورودی به مدل، تعیین تعداد گام‌های زمانی پیشین برای مشاهده حرکت و تعیین تعداد گام‌های زمانی برای پیش‌بینی: در این پژوهش و از میان ویژگی‌های استخراج شده، (مطابق بخش ۴-۳) زاویه حرکت، سرعت و موقعیت به عنوان ویژگی‌های ورودی بررسی می‌شوند. برای افق زمانی پیش‌بینی هم تا ۵ ثانیه بررسی و مقدار رشد خطا اندازه‌گیری می‌شود.

جداسازی داده‌ها: مانند بسیاری از پژوهش‌ها، نسبت ۷۰٪، ۱۵٪ و ۱۵٪ درصد به ترتیب برای داده‌های آموزش، اعتبارسنجی و ارزیابی اعمال می‌شود.

۳-۵-۳-۲- گام دوم: انتخاب نوع مدل

در این پژوهش عملکرد هر ۳ مدل به صورت جداگانه بررسی شده و در بخش نتایج به تفصیل تحلیل می‌شود.

۳-۵-۳- گام سوم: تنظیم پارامترها و شاخص‌ها

پارامترها: تنظیم پارامترها در یک مدل LSTM، یک فرآیند تجربی است که شامل تکرار آموزش مدل به دفعات با مقادیر مختلف برای پارامترها به منظور یافتن بهترین عملکرد است. پارامترهایی مانند Dropout, Units, Learning rate, Batch size و Epochs مقادیر ثابتی برای همه حالات ندارند و مقادیر بهینه آنها بسته به اندازه و پیچیدگی داده‌ها، طول توالی (گام‌های زمانی سابق برای مشاهده) و وظیفه خاص مدل مورد نظر تغییر می‌کند. از آنجا که این عوامل به روش‌های پیچیده و غیرخطی با هم تعامل دارند، جستجوی مقادیر "درست" نیاز به آزمایش‌های سیستماتیک، اعتبارسنجی روی داده‌ها و رصد دقیق فرآیند آموزش و عملکرد نهایی دارد. با تنظیم و اجرای مجدد مدل به صورت مکرر، الگوهایی پدیدار می‌شوند که فرآیند انتخاب را به سمت ترکیبی از پارامترهای متناسب با نیازهای منحصر به فرد مسئله هدایت می‌کنند.

تابع خطا: انتخاب یک تابع خطای مناسب به عملکرد مدل و ماهیت داده‌ها بستگی دارد. در این پژوهش که از نوع مسئله رگرسیونی است از تابع خطای میانگین مربعات (MSE) استفاده شده است.

$$MSE = \frac{1}{N} \left(\sum_{i=1}^N (\hat{y}_i - y_i)^2 \right) \quad (8)$$

در این رابطه، N تعداد کل داده‌ها، \hat{y}_i مقدار پیش‌بینی شده و y_i مقدار واقعی هستند.

الگوریتم بهینه‌سازی: وظیفه تابع خطای کمی‌سازی اختلاف بین پیش‌بینی و مقادیر واقعی است. فرآیند بهینه‌سازی تابع خطا، مقداردهی تکرارشونده پارامترهای یک مدل با هدف کاهش این اختلاف است. در این پژوهش از الگوریتم گرادیان کاهشی تصادفی (SGD) استفاده می‌شود. SGD یک الگوریتم رایج برای آموزش شبکه‌های عمیق است. SGD به جای محاسبه گرادیان تابع خطا با استفاده از کل مجموعه داده‌ها - مانند گرادیان کاهشی استاندارد - پارامترهای مدل را با استفاده از یک نمونه داده و یا یک دسته کوچک از داده‌ها در هر مرحله بروزرسانی می‌کند (رابطه (۹)). این امر باعث می‌شود که برای مجموعه داده‌های بزرگ، سریع‌تر باشد و تصادفی بودن آن به اجتناب از کمینه‌های محلی کمک کند.

$$\theta_{new} = \theta_{old} - \eta \cdot \nabla J(\theta_{old}; x^{(i)}; y^{(i)}) \quad (9)$$

در رابطه اخیر، θ_{old} و θ_{new} به ترتیب پارامترهای بروز شده و پارامترهای قبلی مدل هستند (وزن‌ها و سوگیری). η نرخ یادگیری و $\nabla J(\theta_{old}; x^{(i)}; y^{(i)})$ گرادیان تابع خطای J نسبت به پارامترهای فعلی θ_{old} ، که با استفاده از یک نمونه آموزشی $(x^{(i)}, y^{(i)})$ محاسبه شده است. این جنبه «تصادفی» SGD است، زیرا از تقریبی از گرادیان واقعی (که در کل مجموعه داده‌ها محاسبه می‌شود) استفاده می‌کند.

۳-۵-۴- گام چهارم: آموزش مدل

پس از مقداردهی به کلیه پارامترها مطابق گام‌های پیشین، فرایند یادگیری مدل با استفاده از داده‌های آموزشی و تعداد epoch‌های از پیش تعیین شده، انجام می‌شود. بعد از هر epoch، خطا اندازه‌گیری شده (MAE) تا متوجه روند پیشرفت مدل شویم و همچنین از عدم وقوع بیش یا کم برآزش مدل، اطمینان حاصل نماییم.

۳-۵-۵- گام پنجم: انجام پیش‌بینی و مصورسازی نتایج

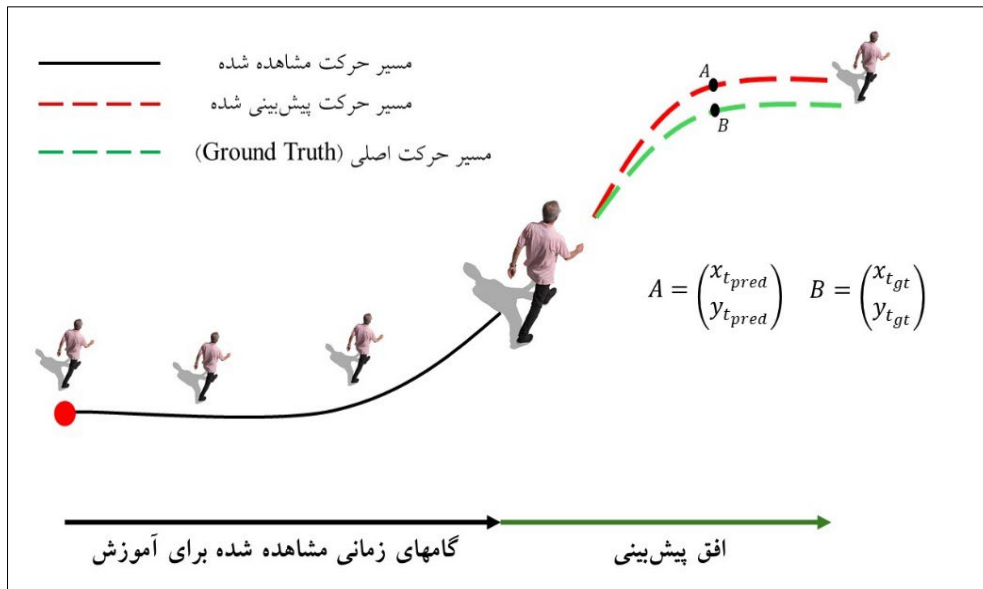
پس از آموزش چندباره مدل و یافتن بهترین مقادیر برای پارامترها، پیش‌بینی بر روی داده‌های دیده نشده (داده‌های ارزیابی) انجام می‌شود و نتایج به شکل‌های مختلف که در بخش ۵ می‌آید، مصور می‌شود.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (10)$$

$$ADE = \frac{1}{(T_{end} - (T_{obs} + 1))} * \sum_{t=T_{obs}+1}^{T_{end}} \sqrt{(\hat{x}_t - x_t)^2 + (\hat{y}_t - y_t)^2} \quad (11)$$

$$FDE = \sqrt{(\hat{x}_{end} - x_{end})^2 + (\hat{y}_{end} - y_{end})^2} \quad (12)$$

- شاخص‌های ارزیابی: پس از انجام آموزش و پیش‌بینی توسط مدل، نیاز به شاخص‌هایی داریم که عملکرد مدل را ارزیابی کند. در این پژوهش در مرحله آموزش از شاخص میانگین خطای مطلق (MAE) و پس از پیش‌بینی مسیر از شاخص‌های میانگین خطای جابجایی^{۲۱} (ADE) و خطای جابجایی نهایی^{۲۲} (FDE) استفاده می‌شود. در روابط (۱۱) و (۱۲)، T_{end} و $T_{obs} + 1$ به ترتیب ابتدای زمان پیش‌بینی و انتهای آن است. همچنین \hat{x}_t و \hat{y}_t مقادیر پیش‌بینی شده برای موقعیت عابر و x_t و y_t مقادیر واقعی آن است.



شکل ۱۱. فرآیند مشاهده مسیر حرکت عابر و پیش‌بینی آن

۴- نتایج و تفسیر آنها

الف) پارامترهای هر مدل

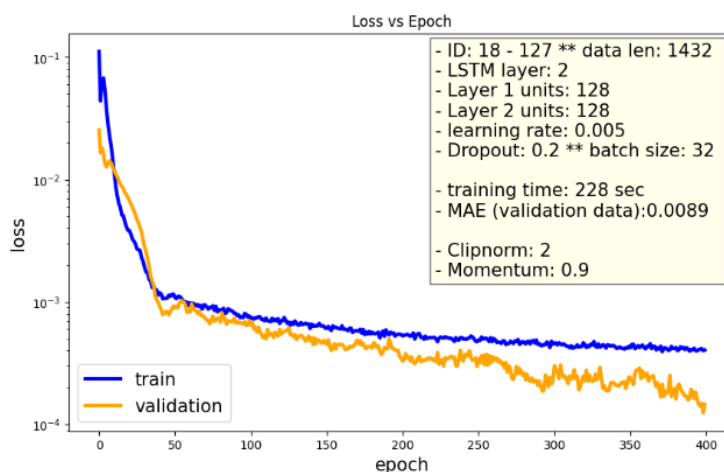
پس از انجام فرایند آموزش به طور جداگانه برای هر مدل به دفعات متعددی و مقادیر مختلفی مختلف برای هر پارامتر و مشاهده نتیجه و ترسیم نمودار تغییرات تابع خطا بر حسب دوره‌های آموزش (شکل ۱۳)، جدول ۴ تهیه شد که در واقع نشان می‌دهد بهترین عملکرد هر مدل با پارامترهای زیر اتفاق می‌افتد. برای آموزش از حرکت عابری با ۱۴۳۲ نمونه و هر نمونه شامل ۵۰ فریم (۲ ثانیه) از حرکت پیشین کاربر می‌باشد. مدل‌ها در ۴۰۰ دور آموزش دیدند.

مطابق جدول ۴، مدل GRU در مقایسه با سایر مدل‌ها، به میانگین خطای مطلق (MAE) کمتری دست یافت. این نتیجه را می‌توان تا حدی به مکانیسم ساده‌تر GRU نسبت داد که تعداد پارامترهای قابل یادگیری را نسبت به LSTM کاهش می‌دهد. علاوه بر این، دروازه‌های بروزرسانی و تنظیم مجدد در GRU سازگاری مؤثرتری با وابستگی‌های زمانی کوتاه‌مدت و میان‌مدت دارد. این نتایج نشان داد که برای مجموعه داده‌هایی با تعداد نمونه نه چندان زیاد و گام‌های زمانی پیشین کم، GRU تعادل قابل قبولی بین پیچیدگی مدل و دقت پیش‌بینی ارائه می‌دهد و در صورت عدم وجود وابستگی‌های بلندمدت، از انواع پیچیده‌تر LSTM عملکرد بهتری دارد. این نتیجه

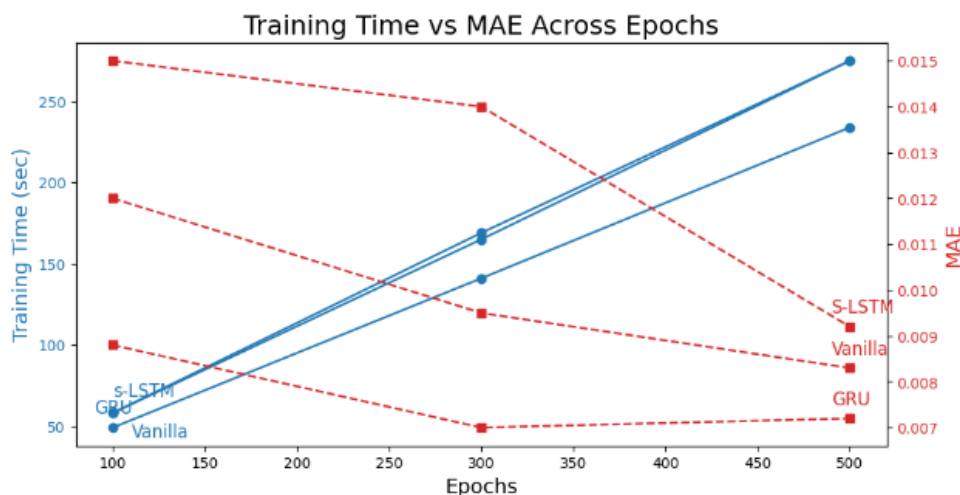
مطابق یافته‌های یانگ و همکاران در سال ۲۰۲۰ است که نشان دادند از نظر سرعت آموزش مدل، GRU برای پردازش مجموعه داده یکسان، ۲۹،۲۹ درصد سریع‌تر از LSTM است؛ و از نظر عملکرد، GRU در مجموعه داده کوچک از LSTM عملکرد بهتری دارد (Yang et al, 2020).

جدول ۳. پارامترهای بهینه هر مدل به تفکیک

مشخصات	GRU	S-LSTM	Vanilla
تعداد لایه‌ها	۲	۲	۱
تعداد واحدها	لایه اول: ۲۵۶ لایه دوم: ۲۵۶	لایه اول: ۱۲۸ لایه دوم: ۱۲۸	۲۵۶
dropout	۰،۲	۰،۲	۰،۲
Learning rate	۰،۰۰۳	۰،۰۰۵	۰،۰۰۶
Batch size	۳۲	۳۲	۳۲
MAE (Validation Data)	۰،۰۰۷۱	۰،۰۰۸۹	۰،۰۰۸۶
زمان آموزش (s)	۲۵۱	۲۳۳	۱۸۸



شکل ۱۲. نمودار تغییرات تابع خطا بر حسب دوره‌های آموزش. ترسیم این نمودار به رصد چگونگی پیشرفت آموزش، شناسایی بیش یا کم برازش مدل و ارزیابی سرعت همگرایی کمک می‌کند. نمودار بالا مربوط به مدل S-LSTM است.



شکل ۱۳. نمودار تغییرات زمان آموزش و شاخص MAE بر حسب دوره‌های آموزش. خطوط خط‌چین برای شاخص MAE (محور قائم سمت راست) و خطوط پیوسته برای زمان آموزش (محور قائم سمت چپ) است.

بیشتر مقادیر این ۲ ویژگی نسبت به ۲ ویژگی دیگر باشد که در یادگیری بلندمدت مدل ایجاد بی‌نظمی می‌کند.

جدول ۴: عملکرد مدل برای ویژگی‌های مختلف

Input Features	Output Features	MAE
X, Y	X, Y	۰/۰۷۸
X, Y, heading, Velocity	X, Y, heading, Velocity	۰/۰۷۷
X, Y, heading, Velocity	heading, Velocity	۰/۰۲
heading, Velocity	heading, Velocity	۰/۰۲۱

ت) مقادیر ADE و FDE

در این بخش مقادیر مختلف شاخص‌های ADE و FDE برای ۳ عابر پیاده متفاوت در ۲ تقاطع متفاوت با هر ۳ مدل محاسبه و مقایسه می‌شود. با بررسی نتایج مشاهده می‌شود که مدل S-LSTM برای افق پیش‌بینی بلندتر عملکرد بهتری دارد. در میان مدت مدل Vanilla عملکرد خوبی دارد و در بازه زمانی کوتاه‌تر، مدل GRU. مدل S-LSTM با لایه‌های چندگانه، توانایی خود در استخراج وابستگی‌های زمانی پیچیده را افزایش می‌دهد و آن را به ویژه برای پیش‌بینی بلندمدت مؤثر می‌سازد. شبکه عمیق‌تر، امکان نمایش سلسله مراتبی الگوها را فراهم می‌کند و در سناریوهایی که مشاهدات گذشته بر مسیرهای

(ب) شاخص MAE و زمان آموزش بر حسب دوره‌های آموزش

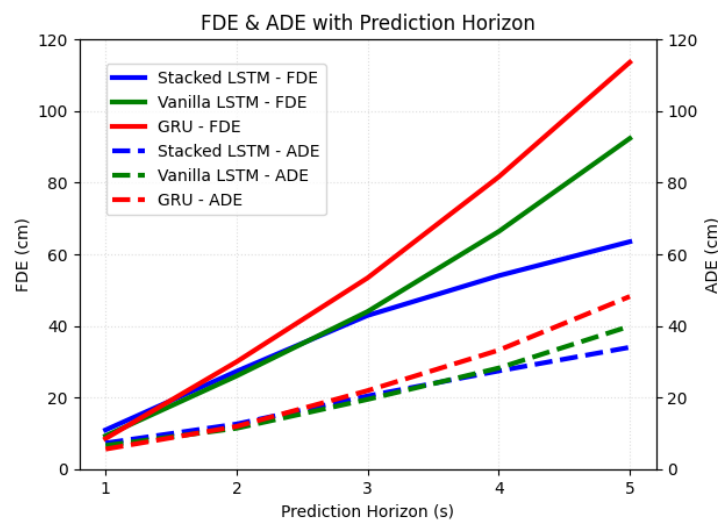
شاخص MAE به عنوان معیاری برای عملکرد مدل تابعی از تعداد دوره‌های آموزش (epoch) است. همچنین زمان صرف شده برای آموزش که به عنوان هزینه آموزش نیز عنوان می‌شود، معیاری برای بصره بودن آموزش مدل است که مستقیماً تابعی از epoch می‌باشد. لذا مطابق شکل ۱۴، تغییرات این دو شاخص بر حسب epoch برای هر ۳ مدل اندازه‌گیری گردید. همانطور که انتظار می‌رفت با افزایش دوره‌های آموزش، زمان آموزش و دقت مدل افزایش می‌یابد (کاهش مقدار MAE). مطابق شکل دیده می‌شود که مدل GRU در حدود ۳۰۰ دور با زمان آموزش حدود ۱۷۰ ثانیه، دارای کمترین مقدار برای شاخص MAE بوده و تاییدی بر نتیجه قسمت قبل است که برای مجموعه داده‌هایی با تعداد نمونه نه چندان زیاد و گام‌های زمانی پیشین کم، GRU تعادل قابل قبولی بین هزینه آموزش مدل و دقت پیش‌بینی ارائه می‌دهد و در صورت عدم وجود وابستگی‌های بلندمدت، از انواع LSTM عملکرد بهتری دارد.

پ) انتخاب ویژگی‌های مناسب برای داده‌های ورودی

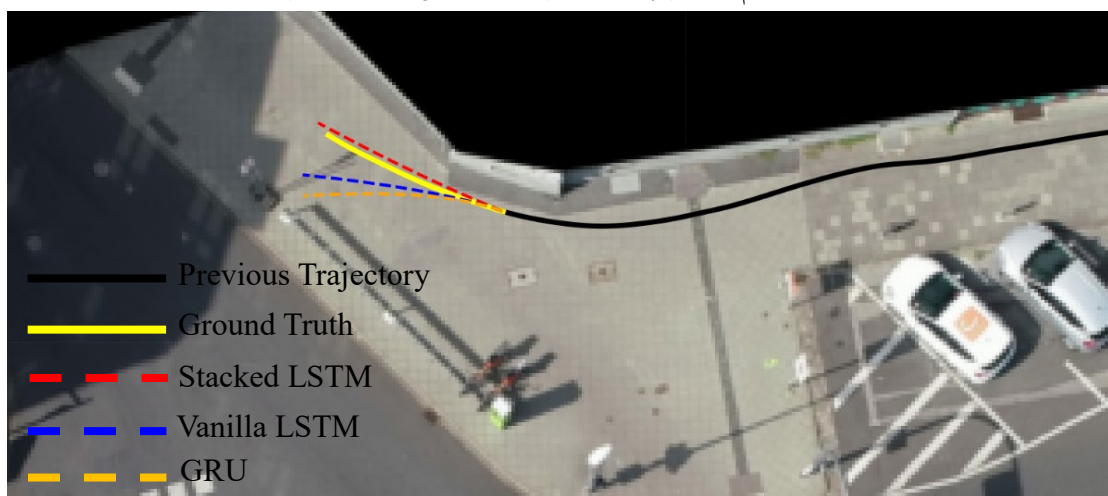
ویژگی‌های ورودی به مدل در حالات مختلف مطابق جدول ۵ به مدل S-LSTM داده و شاخص MAE برداشت گردید. مشاهده می‌شود که ۲ ردیف آخر نسبت به سایر ردیف‌ها عملکرد به مراتب بهتری دارند. علت اینکه در ردیف‌هایی که پیش‌بینی ۲ ویژگی X و Y باعث خطای بالای مدل می‌شود شاید به پراکندگی

در مقایسه با LSTM، سریع تر و کم مصرف تر هستند. با ترکیب دروازه های بروزرسانی و تنظیم مجدد، وابستگی های کوتاه مدت را مدیریت می کنند و آن ها را به ویژه برای پیش بینی های کوتاه مدت، جایی که پاسخگویی به ورودی های اخیر بسیار مهم تر از استخراج الگوهای زمانی عمیق است، مفید می کنند. پارامترهای کمتر، همگرایی سریع تر را ممکن می سازند، اما می توانند اثربخشی را در کارهایی که نیاز به نگهداری طولانی مدت در حافظه دارند، محدود کنند.

آینده تأثیر می گذارند، پیش بینی را بهبود می بخشد. با این حال، افزایش عمق همچنین با هزینه های محاسباتی و چالش پیش برآزش همراه است. مدل Vanilla LSTM تعادلی قابل قبول ارائه می دهد و آن را برای پیش بینی های میان مدت مناسب می سازد. این مدل، وابستگی های ضروری را بدون بار محاسبات زیاد پارامترها ثبت می کند و نتایج پایدار و قابل تفسیری ارائه می دهد. این امر، مدل را به انتخابی ایده آل برای سناریوهایی تبدیل می کند که در آن ها اطلاعات گذشته مفید است اما مدل سازی بلندمدت ضروری نیست. مدل GRU - با یک ساختار ساده شده - مدیریت حافظه سلولی را ساده می کند و



شکل ۱۴. نمودار تغییر شاخص ADE (محور قائم سمت راست و خطوط خطچین در شکل) و FDE (محور قائم سمت چپ و خطوط پیوسته در شکل) بر حسب افق پیش بینی.



شکل ۱۵. مصورسازی پیش بینی ۳ مدل در مسیر حرکت عابر پیاده در افق پیش بینی ۴ ثانیه

۵- نتیجه گیری

بر پیش‌بینی تاثیرگذار باشند. به عنوان پیشنهاداتی در راستای این پژوهش می‌توان گفت که نیاز به داشتن بانک‌های اطلاعاتی کلان و سطح بالا یکی از نیازهای اصلی امروز برای جامعه پژوهشی حمل و نقل ایران در حوزه‌های مختلف از جمله هوش مصنوعی است. وجود این داده‌ها و درک و دریافت انواع الگوها از آنها می‌تواند به تصمیم‌گیری‌های مدیریتی در حوزه حمل و نقل کمک قابل توجهی کند. از طرف دیگر ورود خودروهای خودران به سیستم حمل و نقل ایران، دیر یا زود، اجتناب‌ناپذیر است.

در بسیاری از کشورهای پیشرفته، این فناوری در حال رشد است و ایران نیز باید از همین حالا به سوی آماده‌سازی زیرساخت‌ها حرکت کرده تا بتواند از این تحول بهره‌بردار.

برای آماده شدن در این حوزه ۳ موضوع زیر را می‌توان از اولین گام‌ها برای آمادگی ورود خودران‌ها دانست: ۱. جمع‌آوری داده‌های رانندگی درون شهری ایران و مطالعه بر روی آمادگی ذهنی و فرهنگی جامعه برای پذیرش خودران‌ها ۲. بازنگری در قوانین ترافیکی ۳. توسعه فرهنگ و سواد هوش مصنوعی در صنعت حمل و نقل کشور.

۶- پی‌نوشت‌ها

1. Recurrent Neural Network (RNN)
2. Auto Regressive (AR)
3. Moving Average (MA)
4. Auto Regressive Moving Average (ARMA)
5. Auto Regressive Integrated Moving Average (ARIMA)
6. Support Vector Machine (SVM)
7. Long – Short Term Memory (LSTM)
8. Gated Recurrent Unit (GRU)
9. Attention Mechanism
10. Artificial Neural Network (ANN)
11. Stochastic Gradient Descent (SGD)
12. Back Propagation (BP)
13. Exploding Gradient
14. Vanishing Gradient
15. Input Gate
16. Forget Gate
17. Output Gate
18. Cell State
19. Bias
20. Back Propagation Through Time (BPTT)

در ارزیابی عملکرد معماری‌های مختلف شبکه عصبی بازگشتی برای پیش‌بینی سری‌های زمانی، یافته‌های ما نشان می‌دهد که مدل S-LSTM در پیش‌بینی بلندمدت از هر دو مدل دیگر بهتر عمل می‌کند. معماری عمیق‌تر S-LSTM، آن را قادر می‌سازد تا وابستگی‌های زمانی پیچیده‌تری را ثبت کند، که آن را به‌ویژه برای افق‌های پیش‌بینی طولانی‌مدت مؤثر می‌سازد. برعکس، مدل‌های GRU به دلیل ساختار ساده و همگرایی سریع‌ترشان، عملکرد قوی در پیش‌بینی کوتاه‌مدت نشان می‌دهند، در حالی که Vanilla LSTM با ایجاد تعادل، نتایج قابل اعتمادی را در پیش‌بینی‌های میان‌مدت نشان می‌دهد.

با این حال، توجه به این نکته ضروری است که این نتایج به عواملی مانند شکل و اندازه داده‌ها، ویژگی‌های ورودی، طول توالی و تکنیک‌های پیش‌پردازش بسیار حساس هستند. تعامل این عناصر می‌تواند به طور قابل توجهی بر رفتار مدل تأثیر بگذارد و بر نیاز به طراحی دقیق و تنظیم متناسب با نوع پیش‌بینی تأکید کند. یافته‌های ما تاییدی بر یافته‌های پژوهش کاهوانتری و همکاران (۲۰۲۳) بود نشان دادند نرخ یادگیری و تعداد واحدها در هر لایه از جمله مهمترین پارامترهایی هستند که باید تنظیم شوند و به طور کلی، GRUها در توالی‌های با پیچیدگی کم، عملکرد بهتری نسبت به شبکه‌های LSTM دارند، در حالی که در توالی‌های با پیچیدگی بالا، LSTMها عملکرد بهتری دارند. همچنین در تایید یافته‌های ما، نوشیان و همکاران (۲۰۲۱) نشان دادند که GRU چند لایه مانند مدل ما در این پژوهش از سه مدل دیگر پژوهش یاد شده (vanilla LSTM, stacked LSTM و vanilla GRU) بهتر عمل کرده است.

در طراحی مدل‌ها و پیش‌بینی مسیر حرکت با دیتاست معرفی شده، محدودیت‌هایی در پژوهش وجود داشت از جمله اینکه دیتاست ind برخی عوامل زمینه‌ای مرتبط مانند وجود موانع، شیب و جزئیات زمین، و تعاملات با سایرین و ... را در بر نمی‌گیرد و مدل‌سازی دقیق مسیر را محدود می‌کند، ویژگی‌های موجود ممکن است به طور کامل منعکس‌کننده تأثیر محیطی یا تعاملات اجتماعی نباشند، که به طور قابل توجهی بر حرکت در دنیای واقعی تأثیر می‌گذارند. ویژگی‌هایی مثل سن، جنسیت، صحبت کردن با گوشی، رفتار گروهی، نور و شرایط جوی و امور فوری شخصی (مثل دویدن به قصد رسیدن به اتوبوس) و ... در این دیتاست وجود نداشت که می‌تواند

۷- مراجع

- Salehinejad, H., Sankar, S., Barfett, J., Colak, E., & Valaee, S. (2017). Recent advances in recurrent neural networks. *arXiv preprint arXiv:1801.01078*.
- Shih, S. Y., Sun, F. K., & Lee, H. Y. (2019). Temporal pattern attention for multivariate time series forecasting. *Machine Learning*, 108(8), 1421-1441.
- Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2021). Deep learning for time series forecasting: a survey. *Big data*, 9(1), 3-21.
- Weerakody, P. B., Wong, K. W., Wang, G., & Ela, W. (2021). A review of irregular time series data handling with gated recurrent neural networks. *Neurocomputing*, 441, 161-178.
- Yamak, P. T., Yujian, L., & Gadosey, P. K. (2019, December). A comparison between arima, lstm, and gru for time series forecasting. *In Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, 49-55.
- Zhao, W., Alwidian, S., & Mahmoud, Q. H. (2022). Adversarial training methods for deep learning: A systematic review. *Algorithms*, 15(8), 283.
- Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., & Savarese, S. (2016). Social lstm: Human trajectory prediction in crowded spaces. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 961-971.
- Bartoli, F., Lisanti, G., Ballan, L., & Del Bimbo, A. (2018). Context-aware trajectory prediction. *In 2018 24th International Conference on Pattern Recognition (ICPR)* IEEE, 1941-1946.
- Bighashdel, A., & Dubbelman, G. (2019). A survey on path prediction techniques for vulnerable road users: From traditional to deep-learning approaches. *In 2019 IEEE intelligent Transportation Systems Conference (ITSC)* IEEE.1039-1046.
- Huynh, M., & Alaghband, G. (2019). Trajectory prediction by coupling scene-LSTM with human movement LSTM. *In International Symposium on Visual Computing*, 244-259. Cham: Springer International Publishing.
- Jiang, X., Lin, W., & Liu, J. (2019). A method of pedestrian trajectory prediction based on LSTM. *In Proceedings of the 2019 2nd International Conference on Computational Intelligence and Intelligent Systems*, 79-84.
- Amalou, I., Mouhni, N., & Abdali, A. (2022). Multivariate time series prediction by RNN architectures for energy consumption forecasting. *Energy Reports*, 8, 1084-1091.
- Bock, J., Krajewski, R., Moers, T., Runde, S., Vater, L., & Eckstein, L. (2020, October). The ind dataset: A drone dataset of naturalistic road user trajectories at german intersections. *In 2020 IEEE Intelligent Vehicles Symposium (IV)*, IEEE. 1929-1934.
- Cahuantzi, R., Chen, X., & Güttel, S. (2023, July). A comparison of LSTM and GRU networks for learning symbolic sequences. *In Science and Information Conference*, 771-785. Cham: Springer Nature Switzerland.
- Feng, W., Guan, N., Li, Y., Zhang, X., & Luo, Z. (2017). Audio visual speech recognition with multimodal recurrent neural networks. *In 2017 International Joint Conference on Neural Networks (IJCNN)*, IEEE. 681-688.
- Gruel, W., & Stanford, J. M. (2016). Assessing the long-term effects of autonomous vehicles: a speculative approach. *Transportation Research Procedia*, 13, 18-29.
- Hansson, S. O., Belin, M. Å., & Lundgren, B. (2021). Self-driving vehicles—an ethical overview. *Philosophy & Technology*, 34(4), 1383-1408.
- Hu, J., Wang, X., Zhang, Y., Zhang, D., Zhang, M., & Xue, J. (2020). Time series prediction method based on variant LSTM recurrent neural network. *Neural Processing Letters*, 52(2), 1485-1500.
- Ji, X. (2018, April). The impact of self-driving cars on existing transportation networks. *In AIP Conference Proceedings*, Vol. 1955, No. 1, 040142. AIP Publishing LLC.
- Liang, Y., Wen, H., Nie, Y., Jiang, Y., Jin, M., Song, D. & Wen, Q. (2024). Foundation models for time series analysis: A tutorial and survey. *In Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining*, 6555-6565.
- Liu, P., Yang, R., & Xu, Z. (2019). How safe is safe enough for self-driving vehicles? Risk analysis, 39(2), 315-325.
- Saleh, K., Hossny, M., & Nahavandi, S. (2017, October). Intent prediction of vulnerable road users from motion trajectories using stacked LSTM network. *In 2017 IEEE 20th International Conference on intelligent transportation systems (ITSC)*, IEEE. 327-332.

The International Journal of Robotics Research, 39(8), 895-935.

-Simon, H. (2009). Neural networks and learning machines. *Pearson International Edition*, 282-283.

-Sun, Q., Jankovic, M. V., Bally, L., & Mougiakakou, S. G. (2018, November). Predicting blood glucose with a lstm and bi-lstm based deep neural network. *In 2018 14th symposium on neural networks and applications (NEUREL)*, IEEE. 1-5.

-Kumar, B. P., & Hariharan, K. (2020). Multivariate time series traffic forecast with long short-term memory based deep learning model. *In 2020 International conference on power, instrumentation, control and computing (PICC)*, IEEE. 1-5.

-Nosouhian, S., Nosouhian, F., & Khoshouei, A. K. (2021). A review of recurrent neural network architecture for sequence learning: *Comparison between LSTM and GRU*.

-Rudenko, A., Palmieri, L., Herman, M., Kitani, K. M., Gavrila, D. M., & Arras, K. O. (2020). Human motion trajectory prediction: A survey.

Investigating the Performance of Various Types of Recurrent Neural Networks in Predicting Time Series Data in Transportation; Data Type: Pedestrian Movement Path on the Sidewalk

Ali Edrisi, Associate Professor, Department of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran.

Mohammad Zahedi, M.Sc., Student, Department of Civil Engineering, K.N. Toosi University of Technology, Tehran, Iran.

E-mail: edrisi@kntu.ac.ir

Received: September 2025- Accepted: February 2026

ABSTRACT

In today's world, artificial intelligence has emerged as an undeniably powerful factor in all human affairs. This intelligence is trained and applied using neural network algorithms and big data. Over time, different types of neural networks have been introduced and developed for various applications, one of which is the Recurrent Neural Network (RNN). Due to its architecture and structure, the RNN model demonstrates acceptable performance on time series data—data in which different features of a phenomenon are recorded at fixed time steps. In transportation, there are many cases of time series data, such as the volume of traffic passing through a specific point at fixed intervals; the number, gender, and other characteristics of subway passengers at different times of the day and night; and movement characteristics such as the position, speed, and acceleration of a traffic agent (e.g., a pedestrian) at any given moment on the sidewalk. The last case, which involves the records and history of pedestrian movement characteristics, forms the basis of this research. In this paper, we use three subsets of RNN—namely Vanilla LSTM, Stacked LSTM, and GRU models—to predict pedestrian trajectories. The main goals of this research are, first, to estimate the future position of a pedestrian in order to identify and eliminate hazardous conditions in interactions between pedestrians and autonomous systems (such as delivery robots) on sidewalks, and second, to examine the performance of these three models on time series data. The results show that in the short-term prediction horizon, the GRU model performs better than the other models. However, as the prediction horizon increases and the data become more complex, the Stacked LSTM model outperforms the others.

Keywords: Trajectory Prediction, Time Series Data, Recurrent Neural Network, LSTM, GRU