فصلنامه مهندسی حمل و نقل، دوره ۱۶، شماره ۲، زمستان ۱۴۰۳، ص. ۴۴۷۴-۴۴۶۱

مقاله علمی-پژوهشی

گسترش مدل نوین مبتنی بر شبکه عصبی پیچشی و نظم دهی حذف تصادفی تغییراتی در برآورد شدت تصادفات

علی توکلی کاشانی (مسئول مکاتبات)، استادیار، گروه مهندسی حمل و نقل، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، ایران

E-mail: alitavakoli@iust.ac.ir

امید عبدالحسین پور محجوبیان، دانش آموخته کارشناسی ارشد، گروه مهندسی حمل و نقل، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، ایران

دریافت: ۱۴۰۲/۰۶/۰۸ پذیرش: ۱۴۰۲/۰۷/۱۷

چکیده

مطالعات شدت ترافیک از جهت تلفات و خسارات ناشی از تصادفات جاده ای از اهمیت زیادی برخوردار هستند. مطالعات مختلفی در راستای مدل سازی شدت ترافیک انجام شده که اکثریت این مدل ها از روش های آماری استفاده کرده اند و استفاده از روش های مبتنی بر شبکه های عصبی در این زمینه محدود بوده است. روش های آماری فرضیات مشخصی در رابطه با داده ها و متغیر ها دارند که در صورت تخطی مساله از آنها، استنباط آماری تحت تاثیر قرار خواهد گرفت. این در حالی است که روش های مبتنی بر شبکه های عصبی این فرضیات را نداشته و از این حیث نسبت به روش های آماری برتری دارند. اگرچه که برخی مشکلات در زمینه بکار گیری شبکه های عصبی وجود دارد. یکی از این اشکالات سختی بهینه سازی ابر پارامتر نظم دهی به علت زمان محاسباتی بالا در شبکه های عصبی است. در این راستا پژوهش پیش رو مدلی بر مبنای شبکه عصبی عمیق پیچشی بیزی ارایه کرده است و با استفاده از روش نظم دهی حذف تصادفی تغییراتی، نیاز به روش های جستجوی شبکه ای و اعتبار سنجی متقابل برای یافتن مقدار بهینه ابر پارامتر نظم دهی را از بین برده است. این مدل، روشی ساختار مند برای پیدا کردن ابرپارامتر نظم دهی توامان با آموزش مدل ارایه کرده است. بدین منظور داده های تصادفات جاده های دو خطه دو طرفه جدا نشده کشور ایران مورد استفاده قرار گرفته اند. مدل پیشنهادی با مدل لوجیت ترکیبی و دو مدل دیگر مبتنی بر شبکه عصبی پیچشی بدون نظم دهی و با نظم دهی حذف تصادفی مقایسه شده است. نتایج نشان می‌دهند که عملکرد مدل پیشنهادی در تمامی شاخص های مقایسه به نسبت سایر مدل ها بهتر بوده است.

واژه‌های کلیدی: مدل سازی شدت تصادفات، یادگیری عمیق، شبکه عصبی پیچشی بیزی، نظم دهی حذف تصادفی تغییراتی

۱. مقدمه

به علت افزایش جمعیت شهر ها و گسترششان در بعد فضایی و افزایش ابعاد آنها تقاضا برای سیتم های حمل و نقلی افزایش یافته است. در نتیجه موضوعات مرتبط با حمل و نقل از جمله مساله ایمنی حمل و نقل در مطالعات سال های اخیر به شدت مورد توجه قرار گرفته است. تصادفات در شبکه حمل و نقل جاده ای، از اصلی ترین عوامل تلفات تلقی می‌شوند. این مساله تا جایی است که تصادفات ششمین عامل مرگ و میر با سالانه ۱/۲۴ میلیون کشته و پنجاه هزار مصدوم در دنیا و به عنوان دومین عامل مرگ و میر با سالانه ۱۷۹۹۴ تلفات در ایران هستند. برای مدل سازی رابطه میان شدت تصادفات و عوامل موثر بر آن مانند عوامل مربوط به راننده، عوامل محیطی و وسایل نقلیه، از مجموعه ای از روش ها پیشرفته آماری استفاده شده است. در حالی که استفاده از روش های مبتنی بر شبکه های عصبی و یادگیری عمیق نیاز به بررسی بیشتری دارد. شبکه های عصبی های عصبی در حوزه های متفاتی در حمل و نقل مورد استفاده قرار گرفته اند [Li, Zhu et al. 2020, Zhou, Yang et al. 2020] که می‌توان به مطالعات برآورد داده های سنسور [Fang, Fei et al. 2017] و مطالعات پیش بینی جریان ترافیک [Do, Vu et al. 2019, Ali, Zhu et al. 2022] اشاره کرد، اما استفاده از روش های مبتنی بر یادگیری عمیق در مدل سازی شدت تصادفات محدود تر بوده است. مطالعات متفاوتی برتری عملکردی شبکه های عصبی بر روش های آماری در مدل سازی شدت تصادفات را نشان داده اند [Silva, Andrade et al. 2020, Shiran, Imaninasab et al. 2021]. اگرچه که برخی اشکالات اساسی در بکارگیری و استفاده از شبکه های عصبی عمیق در مدل سازی شدت تصادفات وجود دارد.

یکی از این اشکالات مبحث بیش برازش در شبکه های عصبی است [Bejani and Ghatee 2019, Bejani and Ghatee 2021]. بیش برازش زمانی اتفاق میافتد که مدل فقط داده های واقع در مجموعه آموزشی را یاد میگیرد، در نتیجه عملکردش بر روی داده های آموزش قابل قبول است اما در مواجهه با داده های تست عملکرد ضعیفی نشان خواهد داد. انواع روش های نظم دهی برای جلوگیری از بیش برازش ارایه شده اند، در این بین می توان به نظم دهی L_1 و L_2 و همچنین نظم دهی حذف تصادفی اشاره کرد. اگرچه که این روش ها می توانند تا حد خوبی بیش برازش را کنترل کنند، اما ایراداتی نیز وجود دارد که در ادامه به آنها اشاره شده است:

۱. اول اینکه روش پیدا کردن ابرپارامتر بهینه نظم دهی از طریق عملیات های پی در پی اعتبار سنجی متقابل است. این امر در شبکه های عصبی عمیق که پیچیدگی محاسباتی بالایی دارند میتواند به شدت زمان بر باشد.

۲. دوم اینکه مقدار ابرپارامتر نظم دهی حاصل از این روش لزوما برابر با مقدار بهینه آن نیست، چرا که جستجو برای این ابرپارمتر در فضای گسسته انجام می‌شود که لزوما این مقدار بهینه را شامل نمی‌شود.

در نتیجه، پژوهش پیش رو با ارایه یک مدل نوین مبتنی بر شبکه عصبی عمیق بیزی و با استفاده از نظم دهی حذف تصادفی تغییراتی سعی در بهبود عملکرد نظم دهی، چه از منظر عملکرد کلی در کاهش بیش برازش و چه از منظر کاهش عملیات های اعتبار سنجی متقابل دارد. همچنین مدل پیشنهادی با استفاده از روش تغییر متغیر محلی برای برآورد گرادیان در شبکه عصبی عمیق بیزی، عملکرد کلی و دقت مدل سازی را نیز افزایش می‌دهد.

از این حیث، پژوهش پیش رو در ادامه مطالعات حاضر در زمینه ایمنی ترافیک در جهت به کار گیری شبکه های عصبی برای مدل سازی شدت تصادفات است و سعی دارد تا: (۱) روشی ساختار مند برای بهینه سازی ابرپارامترنظم دهی ارایه کند؛ (۲) عملکرد کلی شبکه های عصبی را با استفاده از شبکه های عصبی بیزی افزایش دهد؛ (۳) مقایسه ای از عملکرد مدل پیشنهادی با

سایر مدل های شبکه های عصبی متداول در ادبیات پژوهشی و همچنین سایر الگوریتم های متداول یادگیری ماشین ارایه کند.

۲. مرور ادبیات

۲-۱ روش های آماری در مدل سازی شدت تصادفات

در مطالعات مرتبط با مدل سازی شدت تصادفات، روش های آماری بیشترین استفاده را داشته‌اند. از آنجا که شدت تصادفات در سطوح گسته دسته بندی می‌شوند، استفاده از مدل های گسسته مورد بررسی قرار گرفته است. به عنوان نمونه از برخی از این روش ها می توان به مدل های لوجیت و پروبیت دوگانه [Do, Vu et al. 2019, Fanyu, Sze et al. 2021, Zubaidi, Obaid et al. 2021] و چند گانه [Vajari, Aghabayk et al. 2020, Park and Park 2022] اشاره کرد. برای بررسی اثرات ترتیبی در شدت تصادفات، ناهمگونی مشاهده نشده، همبستگی بین تصادفات و درونزایی روش های پیشرفته آماری و اقتصاد سنجی از جمله مدل های سلسله مراتبی [Rezapour, Wulff et al. 2020]، مدل لوجیت آشیانه ای [Patil, Geedipally et al. 2012, Islam, Kelarestaghi et al. 2019] و همچنین مدل های با پارامتر تصادفی [Zeng, Wen et al. 2017, Hou, Huo et al. 2022] اشاره کرد. این مدل ها پیش فرض هایی مانند روابط خطی مشخص بین متغیر مستقل و متغیر های توضیح، توابع اتصال و همچنین فرض توزیع مشخص بر روی داده های تصادفات دارند. در صورتی ارضا نشدن این مفروضات، استنباط آماری تحت تاثیر قرار خواهد گرفت [Silva, Andrade et al. 2020, Hou, Huo et al. 2022].

۲-۲ شبکه های عصبی در مدل سازی شدت تصادفات

در برخی از مطالعات، شبکه های عصبی برای مدل سازی شدت تصادفات مورد استفاده قرار گرفته‌اند. برخی از مزایای استفاده از این الگوریتم ها به اختصار مورد بحث قرار می‌گیرد. شبکه های عصبی پیش فرض های مدل ها آماری را ندارند، همچنین این شبکه ها می‌توانند روابط غیر خطی را نیز مدل سازی کنند. برخی از پژوهش هایی که از شبکه های عصبی برای مدل سازی شدت تصادفات استفاده کرده اند در ادامه مورد بررسی قرار گرفته‌اند. عدالوهاب و عبدلعاطی در مطالعه ای با استفاده از شبکه پرسپترون چند لایه شدت تصادفات را مدل سازی کردند. نتایج نشان داد که عملکرد مدل مبتنی بر شبکه عصبی به نسبت روش آماری مدل لوجیت ترتیبی بهتر بوده است [Abdelwahab and Abdel-Aty 2001]. همچنین در سایر مطالعاتی که از شبکه های پرپترپن چند لایه استفاده کرده اند، نتایج مشابهی گزارش شده است. دلن و همکاران با استفاده از چند شبکه پرسپترون چند لایه به شناسایی عوامل موثر در شدت تصادفات پرداختند [Delen, Sharda et al. 2006]. همچنین چیمبا و ساندو در مقایسه عملکرد مدل پرسپرتون چند لایه با روش های آماری لوجیت و پروبیت ترتیبی دقت بالاتری را برای مدل های شبکه عصبی گزارش کردند [Chimba and Sando 2009]. همچنین در مطالعات [Zeng and Huang 2014, Rahim and Hassan 2021] نیز دقت عملکرد مدل های مبتنی بر شبکه های عصبی بالاتر از روش های آماری گزارش شده‌اند. شبکه های عصبی عمیق نیز در برخی مطالعات در مدل سازی شدت تصادفات مورد استفاده قرار گرفته‌اند [Rezapour, Nazneen et al. 2020, Rahim and Hassan 2021]. برخی از این شبکه های شامل شبکه های عصبی پیچشی [Yu, Li et al. 2021, Haghshenas, Guido et al. 2023] و شبکه های عصبی بازگشتی [Rezapour, Nazneen et al. 2020, Zuo, Qian et al. 2023] و همچنین شبکه های پیچشی-بازگشتی [Li, Abdel-Aty et al. 2020, Hu, Zhou et al. 2022] می باشند. اگرچه عملکرد این شبکه ها در مدل سازی شدت تصادفات بسیار امیدوار کننده بوده، استفاده

از این شبکه ها بسیار محدود بوده است [Sameen and Pradhan 2017, Zeng, Gu et al. 2019].

۳. روش شناسی

۳-۱ استنباط بیزی تغییراتی

استباط بیزی، مکتبی در مبحث استنباط آماری است که در برابر مکتب فراوانی گرایانه مطرح شده‌است. برآورد پارامتر های مدل، با توجه به مشاهدات و داده ها و با توجه به اطلاعات اولیه محقق در رابطه پارامتر ها صورت می‌گیرد. برای بردار پارامتر های مدل $\boldsymbol{w}$ و با توجه زوج داده های $\boldsymbol{\mathcal{D}} = \{(\boldsymbol{x_i}, \boldsymbol{y_i})\}_{\boldsymbol{i=1}}^{\boldsymbol{n}}$ که در آن برای هر زوج داده داریم $\boldsymbol{y_i} \in \mathbb{R}$ و $\boldsymbol{x_i} \in \mathbb{R}^{\boldsymbol{p}}$، با مشروط کردن توزیع احتمالی پارامتر های مدل بر داده ها، رابطه استنباط بیزی را به صورت زیر می‌نویسیم:

$$\boldsymbol{p(w|y,X)} = \frac{\boldsymbol{p(y|w,\ X)p(w)}}{\int \boldsymbol{p(y|w,\ X)p(w)dw}} \qquad (۱)$$

از آنجایی که در اکثر موارد عبارت بالا فرم بسته‌ای ندارد، در مطالعات از روش های عددی کمک گرفته شده است.

استنباط بیزی تغییراتی یکی از روش های عددی ارایه شده برای حل این مساله است. روش استباط تغییراتی نسبت به روش MCMC سریع تر و مقیاس پذیر تر **است** [Blei, Kucukelbir et al. 2017] و از این حیث برای برآورد پارامتر های شبکه های عصبی عمیق مورد استفاده قرار می‌گیرد. برای توزیع جایگزین $\boldsymbol{q_\Phi(w)}$ که توزیعی با پارامتر های $\boldsymbol{\Phi}$ است، برآورد تغییراتی این توزیع جایگزین را با کمینه کردن فاصله این توزیع از توزیع پسین اصلی که توسط معیار واگرایی کولبک-لایبلر اندازه گیری می‌شود، بهینه می‌کند. برای توزیع حاشیه ای مشاهدات، با مشروط کردن به صورت ضمنی بر ماتریس متغیر های توضیح $\boldsymbol{X}$، می‌نویسیم:

$$\begin{aligned}\boldsymbol{LN\ p(y)} &= \boldsymbol{LN}\ \frac{\boldsymbol{p(y,w)}}{\boldsymbol{q_\Phi(w)}} \\ &\quad - \boldsymbol{LN}\ \frac{\boldsymbol{p(w|y)}}{\boldsymbol{q_\Phi(w)}} =\end{aligned} \qquad (۲)$$

$$\begin{aligned}&\boldsymbol{LN\ p(y,w) - LN\ q_\Phi(w)} \\ &\quad - \boldsymbol{LN}\ \frac{\boldsymbol{p(w|y)}}{\boldsymbol{q_\Phi(w)}}\end{aligned}$$

که با محاسبه امید ریاضی در دو طرف رابطه داریم:

$$\begin{aligned}&\boldsymbol{LN\ p(y)} \\ &= \int \boldsymbol{q_\Phi(w) LN\ p(w,y)dw} \\ &- \int \boldsymbol{q_\Phi(w) LN\ q_\Phi(w)dw} \\ &+ \int \boldsymbol{q_\Phi(w) LN \frac{q_\Phi(w)}{p(w|y)} dw}\end{aligned} \qquad (۳)$$

و به اختصار می‌نویسیم:

$$\boldsymbol{LN\ p(y) = \mathcal{L} + D_{KL}[q||p]} \qquad (۴)$$

که در آن $\boldsymbol{\mathcal{L}}$ کران پایین لگاریتم توزیع حاشیه ای مشاهدات یا گواه مشاهدات است و $\boldsymbol{D_{KL}[q||p]}$ واگرایی کولبک-لیبلر بین توزیع جایگزین و توزیع پسین پارامتر های مدل است.

از آنجایی که واگرایی کولبک-لیبلر به تعریف همواره دارای مقداری مثبت است، و از آنجایی که گواه مدل، مستقل از پارامتر های توزیع جایگرین می‌باشد و به نسبت تغییرات در این پارامتر ها دارای مقدار ثابتی است، بیشینه کردن کران پایین گواه، متناظر با کمینه کردن واگرایی کولبک-لیبلر توزیع جایگرین و توزیع پسین می‌باشد. در نتیجه با توجه به توضیحات، قرینه کران پایین گواه مدل می تواند به عنوان تابع هزینه در روند آموزش مدل استفاده شود. در نتیجه می نویسیم:

$$\begin{aligned}\boldsymbol{\mathcal{L}} &= -\boldsymbol{D_{KL}[q_\Phi(w)||p(w)]} \\ &\quad + \boldsymbol{\mathbb{E}_{q_\Phi(w)}[p(y|w)]}\end{aligned} \qquad (۵)$$

بیشینه کردن مقدار کران پایین گواه مشاهدات متناظر با کمینه کردن قرینه این مقدار به عنوان تابع هزینه در فرآیند آموزش مدل است.

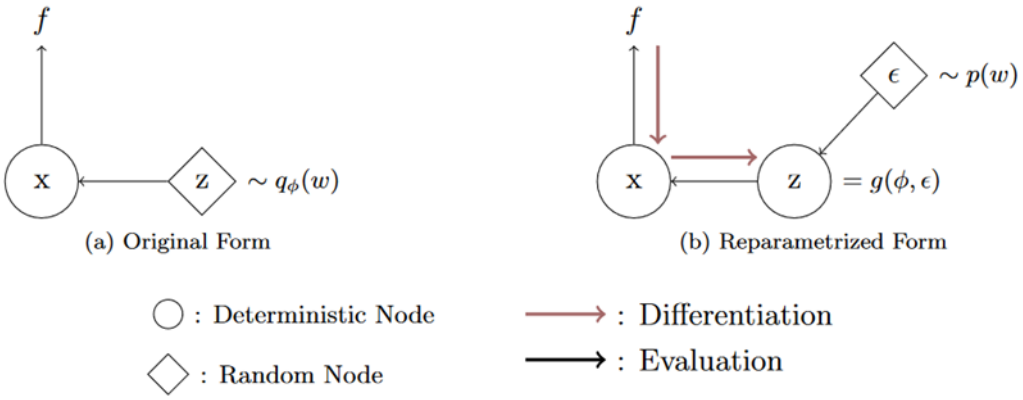
۳-۲ تغییر متغیر محلی

برای عملیات پس انتشار با توجه به تابع هزینه رابطه ۵ داریم:

$$\begin{aligned}\boldsymbol{\nabla_\Phi \mathcal{L}} &= \boldsymbol{\nabla_\Phi \mathbb{E}_{q_\Phi(w)}[LN\ p(y,w)} \\ &\quad \boldsymbol{- LN\ q_\Phi(w)]} \\ &\neq \boldsymbol{\mathbb{E}_{q_\Phi(w)}[\nabla_\Phi (LN\ p(y,w)} \\ &\quad \boldsymbol{- LN\ q_\Phi(w))]}\end{aligned} \qquad (۶)$$

این رابطه به این معناست که به علت آن که عملیات نمونه گیری مشتق پذیر نیست، پس انتشار به صورت مستقیم بر روی نودهای تصادفی گراف عملیاتی شبکه ممکن نیست. تصویر یک بیانگر این موضوع است. در نتیجه برای عملیات پس انتشار با ارایه یک تغییر متغیر مشتق پذیر داریم:

$$\boldsymbol{w = g(\Phi, \varepsilon)}, \qquad \boldsymbol{\varepsilon \sim p(\varepsilon)} \tag{۷}$$



شکل ۱. تغییر متغیر و عملیات پس انتشار. در شکل سمت راست عامل تصادفی برونی سازی شده و در نتیجه عملیات پس انتشار ممکن است.

۳-۳ نزدیک ترین همسایگی

نظم دهی حذف تصادفی بر مبنای ضرب یک سیگنال اختلالی ناخواسته بر روی لایه های مدل کار می‌کند. اگر عناصر ماتریس اختلال را از یک توزیع گاوسی با پارامتر های $\boldsymbol{\mathcal{N}(1, \alpha)}$ انتخاب کنیم، برای یک لایه با ورودی $\boldsymbol{A}$ و وزن های قطعی $\boldsymbol{\theta}$ خواهیم داشت:

$$\begin{aligned} \boldsymbol{q_\Phi(b_{mj}|A) = \mathcal{N}(\psi_{mj}, \xi_{mj}), \quad \psi_{mj}} \\ \boldsymbol{= \sum_{n=1}^{N^{l-1}} a_{mn}\theta_{nj}, \quad \xi_{mj}} \\ \boldsymbol{= \alpha \sum_{n=1}^{N^{l-1}} a_{mn}^2 \theta_{nj}^2} \end{aligned} \tag{۸}$$

که دقیقا متناظر با در نظر گرفتن توزیع جایگزین با پارامتر های $\boldsymbol{w_{ij} = \mathcal{N}(w_{ij}|\theta_{ij}, \alpha\theta_{ij}^2)}$ است. در نتیجه قرار دادن این توزیع بر روی وزن ها شبکه و انجام عملیات پس انتشار، متناظر با نظم دهی تصادفی با پارامتر $\boldsymbol{\alpha}$ است. دقت شود که مقدار بهینه پارامتر نظم دهی در فرآیند آموزش مدل بدست آمده و نیازی به عملیات جستجوی شبکه ای و اعتبار سنجی متقابل نیست.

۴-۳ مدل پیشنهادی

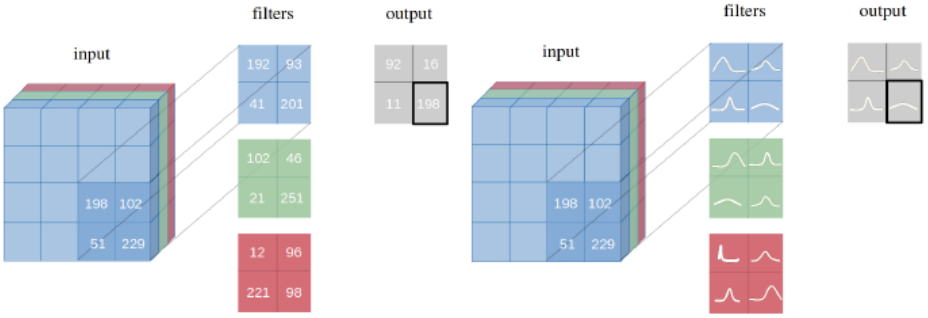
مدل پیشنهادی در پژوهش حاضر، یک شبکه عصبی عمیق پیچشی بیزی است. در این شبکه از استباط بیزی برای برآورد وزن ها استفاده شده. روش های تغییر متغیر محلی به منظور بهبود عملکرد کلی مدل مورد استفاده قرار گرفته و همچنین روش نظم دهی حذف تصادفی تغییراتی برای جلوگیری از بیش برازش مورد استفاده قرار گرفته اند. برای یک فیلتر در شبکه عصبی پیچشی، با توجه به مطالب توضیح داده شده از توزیع زیر استفاده شده است:

$$\begin{aligned} &\boldsymbol{q_\phi(w_{ijhw})} \\ &\boldsymbol{= \mathcal{N}(w_{ijhw}|\theta_{ijhw}, \alpha_{ijhw}\theta_{ijhw}^2)} \end{aligned} \tag{۹}$$

که در آن $\boldsymbol{i}$ لایه ورودی، $\boldsymbol{j}$ لایه خروجی و $\boldsymbol{h}$ و $\boldsymbol{w}$ سطر و ستون فیلتر مورد نظر هستند هستند. شکل دو مقایسه ای از شبکه عصبی پیچشی بیزی و شبکه عصبی پیچشی معمول را نشان می‌دهد. همچنین فرآیند برآورد وزن های شبکه از طریق اعمال دو عملیات

$$\boldsymbol{B_j = A_i * \theta_{ij} + \zeta_j \circ \sqrt{A_i^2(\alpha_{ij} \circ \theta_{ij}^2)}} \tag{۱۰}$$

که در آن * عملیات پیچش و ∘ عملیات ضرب هادامارد است.



شکل ۲. شبکه عصبی پیچشی بیزی (راست) و شبکه عصبی پیچشی معمولی (چپ)

۳-۵ داده های مورد استفاده

در این پژوهش داده های تصادفات جاده های دوخطه دوطرفه جدا نشده بین سال های ۹۴ تا ۹۸، در کشور ایران استفاده شد. این داده ها از گزارشات تصادفات پلیس استخراج شده است. همچنین سایر اطلاعات از سازمان راهداری و همچنین با استفاده از سیستم های اطلاعات مکانی به دست آمده اند. در این داده ها متغیر های مربوط به ویژگی های راننده در کنار مجموعه ای از متغیر های که در مطالعات پیشین موثیر تشخیص داده شده اند قرار گرفته است. متغیر های موجود در این داده در جدول یک گردآوری شده‌اند.شدت تصادفات در سه سطح خسارتی، جرحی و فوتی ارایه شده است. با توجه به تعداد کم مشاهدات تصادفات فوتی و جرحی به نسبت تصادفات خسارتی، تصادفات این دو دسته با هم جمع شدند و یک سطح در نظر گرفته شده‌اند.

ابتدا عملیات پیش پرداز برروی این داده ها انجام شد. داده هایی که مقادیر نامشخص و ثبت نشده داشتند حذف شدند. پس از حذف داده های ثبت نشده، تعداد ۴۱۱۲۰ داده باقی ماند. برای تطبیق بهتر با معماری مدل های شبکه عصبی پیچشی، متغیر های طبقه ای به صورت متغیر ساختگی درآمدند و همچنین با توجه به مطالعات پیشین، داده ها به دو مجموعه آموزش و تست با نسبت ۲۰:۸۰ تقسیم شدند. در ادامه بر روی متغیر های پیوسته عملیات استاندار سازی هم برای داده ای آموزش و هم برای داده های تست و به صورت جداگانه انجام شد.

۴. مدل سازی و نتایج

در این قسمت، نتایج حاصل از مدل های توضیح داده شده را گزارش کرده‌ایم. لازم به ذکر است پیاده سازی تمامی مدل ها با استفاده از زبان برنامه نویسی پایتون انجام شده است. برای مدل لوجیت دوگانه از کتابخانه Sci-kit Learn و برای پیاده سازی مدل های شبکه عصبی از کتابخانه Keras بر روی TensorFlow استفاده کرده‌ایم. مدل ارایه شده با استفاده از برنامه نویسی شی گرا و با تولید لایه های شخصی در Keras انجام شده‌است. در این قسمت عملکرد ۴ مدل با هم مقایسه شده است:

۱. مدل لوجیت دوگانه ترکیبی

۲. سه مدل مبتنی بر شبکه عصبی پیچشی، بدون نظم دهی، با نظم دهی حذف تصادفی و با نظم دهی حذف تصادفی تغییراتی

۴-۱ مشخصات مدل

جدول ۱. متغیر های مورد استفاده

کد متغیر	نام متغیر	توصیف متغیر
Y	شدت تصادف	۰=خسارتی ۱=جرحی فوتی
X_1	سن راننده مقصر	
X_2	جنسیت راننده مقصر	۰=مرد ۱=زن
X_3	تحصیلات راننده مقصر	۰=بیسواد ۱=ابتدایی ۲=دیپلم ۳=لیسانس و بالاتر
X_4	روشنایی	۰=روز ۱=شب
X_5	هوا	۰=صاف ۱=ابری/بارانی ۲=برفی
X_6	نوع منطقه	۰=هموار ۱=کوهستانی
X_7	شرایط سطح راه	۰=تر ۱=خشک
X_8	هندسه محل	۰= پیچ/سربالایی سرپایینی ۱=پیچ مسطح ۲=مستقیم سربالایی/سرپایینی ۳=مستقیم مسطح
X_9	نوع شانه راه	۰=شانه آسفالت ۱=شانه خاکی ۲=شانه ندارد
X_{10}	طول راه	
X_{11}	عرض راه	
X_{12}	سرعت مجاز	
X_{13}	ADT	اطلاعات برداشت شده از سازمان راه داری
X_{14}	HVP	نسبت تردد وسایل نقلیه سنگین به کل تردد سالانه
X_{15}	تعداد رمپ	
X_{16}	TRD	نسبت تعداد دسترسی های هر راه در دوطرف به کل طول راه

برای یافتن بهترین معماری برای مدل پیشنهادی، از آنجایی که برای تعیین این ابرپارامتر ها روش مشخصی موجود نیست، با استناد به پژوهش های پیشین و عملیات های آزمایشی مختلف به مقادیر زیر رسیدیم. مقدار طول ۳ برای طول فیلتر، طول یک برای قدم و همچنین طول ۲ و طول قدم دو در لایه های تجمیعی. و همچنین در مجموع ۳ لایه پیچشی به ترتیب با تعداد ۱۶، ۳۲ و ۶۴ فیلتر.

در آزمایشات مقادیر ۴ لایه و ۵ لایه هم مورد بررسی قرار گرفت ولی عملکرد مدل کاهش یافت که علت آنرا می بایست در بهینه سازی نابالغ جست. جدول ۲ خلاصه ای از اطلاعات معماری مدل پیشنهادی را ارایه می‌کند. همچنین همین عملیات برای سایر مدل های مبتنی بر شبکه عصبی پیچشی انجام شد که اطلاعات متناظر با آن در جدول ۳ گردآوری شده. همچنین برای یافتن مقدار ابرپارامتر نظم دهی از روش جستجوی شبکه ای در فضای ۰/۳ تا ۰/۷ استفاده شده است.

۴-۲ نتایج

در این مرحله داده های مشروح بر روی هریک از مدل ها بارگذاری شده‌اند. در ادامه عملکرد هر یک از این مدل ها بررسی شده است. ابتدا عملکرد مدل لوجیت دوگانه ترکیبی بحث شده است. مدل لوجیت دوگانه ترکیبی همانطور که توضیح داده شد یکی از متداول ترین مدل ها در مطالعات شدت تصادفات است. مدل لوجیت ترکیبی، به علت داشتن پارامتر های تصادفی، به نسبت مدل لوجیت ساده انعطاف پذیر تراست. و توانایی در نظر گرفتن ناهمگونی مشاهده نشده در داده ها را دارد و در نتیجه عملکرد بهتری به نسبت مدل لوجیت ساده داراست. مدل بعدی مورد مطالعه مدلی مبتنی بر شبکه عصبی پیچشی است. در این

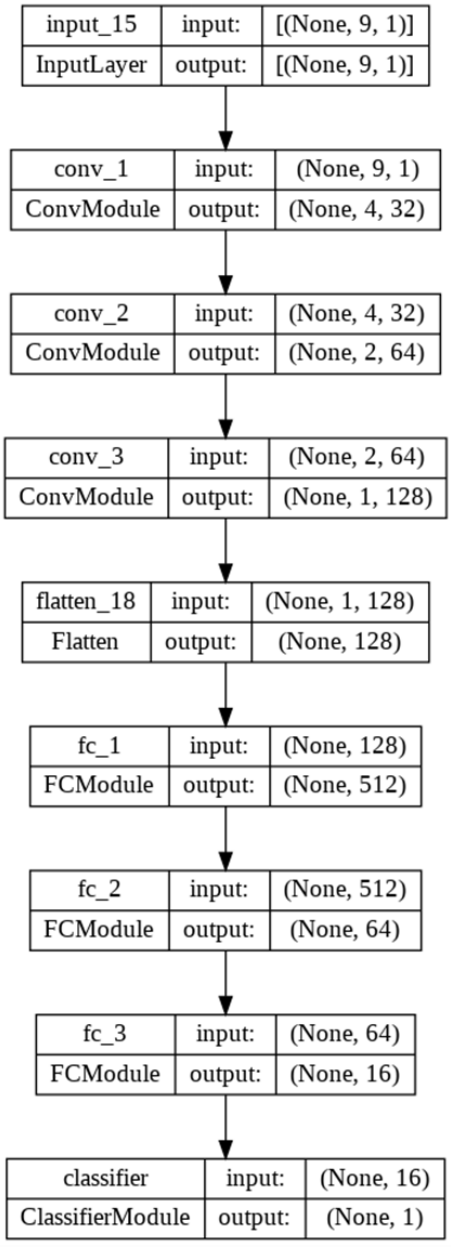
مدل از هیچ روش نظم دهی استفاده نشده است. همچنین یک مدل دیگر میبتنی بر شبکه عصبی پیچشی است. در این مدل از هیچ روش نظم دهی استفاده نشده است. همچنین یک مدل دیگر میبتنی بر شبکه عصبی پیچشی با نظم دهی حذف تصادفی نیز گسترش یافته است. ساختار این دو مدل در جدول ۳ ارایه شده است. همچنین در آخر مدل پیشنهادی با ساختارارایه شده در جدول ۴ مورد بررسی قرار گرفته است. جدول چهار، عملکرد مدل ها را مقایسه کرده است.

جدول ۲. ساختار مدل پیشنهادی

تعداد پارامتر	ابعاد خروجی	لایه
۰	(۹ ،۱)	ورودی
۳۵۲	(۴ ،۳۲)	لایه پیچشی ۱
۱۲۶۰۸	(۲ ،۶۴)	لایه پیچشی ۲
۴۹۷۹۲	(۱ ،۱۲۸)	لایه پیچشی ۳
۱۳۳۶۳۲	۱۲۸	لایه اتصال ۱
۶۵۸۵۶	۵۱۲	لایه اتصال ۲
۲۱۲۸	۶۴	لایه اتصال ۳
۳۳	۱	لایه طبقه بندی
۲۶۴،۴۰۱		تعداد کل پارامتر ها
۲۶۲،۷۶۹		پارامتر های آموزش پذیر
۱،۶۳۲		پارامتر های آموزش ناپذیر

جدول ۳. ساختار مدل شبکه عصبی پیچشی

تعداد پارامتر	ابعاد خروجی	لایه
۰	(۹ ،۱)	ورودی
۱۷۶	(۴ ،۳۲)	لایه پیچشی ۱
۶۰۴	(۲ ،۶۴)	لایه پیچشی ۲
۲۴۸۹۶	(۱ ،۱۲۸)	لایه پیچشی ۳
۶۶،۸۱۶	۱۲۸	لایه اتصال ۱
۳۲،۹۲۸	۵۱۲	لایه اتصال ۲
۱،۰۶۴	۶۴	لایه اتصال ۳
۳۳	۱	لایه طبقه بندی
۱۳۲،۲۱۷		تعداد کل پارامتر ها
۱۳۱،۳۶۸		پارامتر های آموزش پذیر
۸۴۹		پارامتر های آموزش ناپذیر



شکل ۳. شماتیک مدل پیشنهادی

جدول ۴. مقایسه عملکرد مدل ها

LRT-CNN-VD	CNN w/Dropout	CNN	ML

	آموزش	تست	آموزش	تست	آموزش	تست	آموزش	تست
Percision	۰/۵۴	۰/۵۳	۰/۵۹	۰/۶۲	۰/۷۰	۰/۷۱	۰/۷۲	**۰/۷۵**
Recall	۰/۲۹	۰/۲۷	۰/۵۹	۰/۵۰	۰/۴۲	۰/۴۳	۰/۴۶	**۰/۴۹**
F1 Score	۰/۳۸	۰/۳۶	۰/۶۵	۰/۵۵	۰/۵۳	۰/۵۴	۰/۵۴	**۰/۵۹**
Accuracy	۰/۵۸	۰/۵۷	۰/۷۳	۰/۶۳	۰/۶۷	۰/۶۷	۰/۷۱	**۰/۷۲**

در جدول چهار، ML مدل لوجیت ترکیبی، CNN مدل شبکه عصبی پیچشی بدون نظم دهی حذف تصادفی، CNN w/Dropout مدل شبکه عصبی پیچشی با نظم دهی حذف تصادفی و LRT-CNN-VD مدل پیشنهادی است.

همانطور که در جدول ۴ مشخص است، از حیث شاخص عملکرد مدل پیشنهادی در تمامی شاهص ها به نسبت سایر مدل ها بهتر است. از مقایسه شاخص های مختلف، می توان نتیجه گرفت که روش نظم دهی حذف تصادفی عملکرد مدل را بهبود بخشیده است. از مقایسه مقدار شاخص های مختلف بر روی داده های تست این امر مشخص می‌شود. به طور مثال دقت مدل شبکه عصبی پیچشی با حذف تصادفی بر روی داده های تست برابر با ۶۷ درصد است که این مقدار برای مدل شبکه عصبی پیچشی بدون نظم دهی برابر با ۶۳ درصد است. این در حالی است که عملکرد مدل شبکه عصبی پیچشی بدون حذف تصادفی بر روی داده های آموزش برابر با ۷۳ درصد است که بیشترین مقدار را بین تمام مدل ها دارد. در نتیجه اثر بیش برازش مشخص می‌شود. در مقایسه بعدی بین عملکرد مدل پیشنهادی با مدل حذف تصادفی ساده، مشخص می‌شود که عملکرد مدل پیشنهادی به صورت قابل توجهی بهتر است. دلیل این امر یافتن بهینه ابرپارامتر نظم دهی در مدل پیشنهادی است، چرا که فضای جستجو برای این ابر پارامتر فضایی پیوسته است. در عوض در روش نظم دهی حذف تصادفی فضای جستجو فضایی گسسته و محدود در نظر گرفته شده است. همچنین این در حالی است که مدل پیشنهادی نیازی به عملیات های اعتبار سنجی و جستجوی شبکه ای ندارد و از این حیث نیز فرآیند آموزش مدل را تسریع می‌کند.

نکته دیگر حایز اهمیت، عملکرد بهتر مدل هایی که از روش نظم دهی حذف تصادفی استفاده کرده اند، بر روی داده های تست به نسبت داده های آموزش است. علت این امر را نیز، تفاوت نحوه عملکرد روش نظم دهی حذف تصادفی در هنگام آموزش و استنباط است. روش نظم دهی حذف تصادفی تنها در هنگام آموزش عمل می‌کند و شبکه در زمان استنباط، بدون نویز حاصل از نظم دهی انجام می‌شود و در نتیجه، شبکه بر روی داده های تست بهتر است. همچنین در جدول های ۵، ۶، ۷ ماتریس آشفتگی برای هریک از این مدل ها ارایه شده است.

لازم به ذکر است که برای افزایش ثبات نتایج گزارش شده، برای هر مدل ۱۰ بار فرآیند آموزش و تست انجام شده و در نهایت میانگین نتایج در جدول ۴ گزارش شده است.

جدول ۵. ماتریس آشفتگی مدل لوجیت ترکیبی

		برآورد		
		جرحی/ فوتی		خسارتی
مشاهده	کل	۸۲۲۴	۱۸۲۴	۶۴۰۰
	جرحی/ فوتی	۳۵۱۱	۹۶۳	۲۵۴۸
	خسارتی	۴۷۱۳	۸۶۱	۳۸۵۲

جدول ۶. ماتریس آشفتگی مدل شبکه عصبی پیچشی با نظم دهی حذف تصادفی

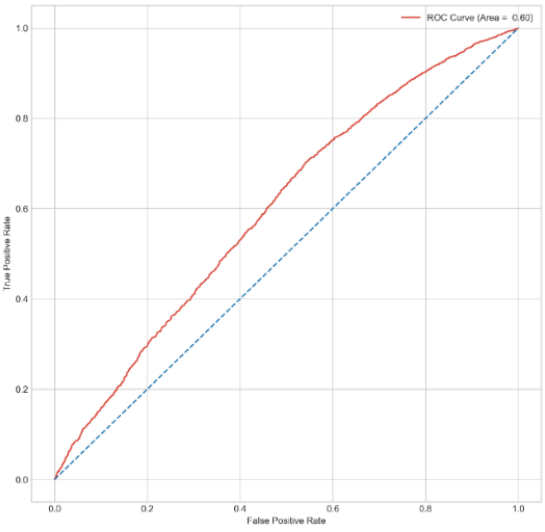
		برآورد		
			جرحی/ فوتی	خسارتی
مشاهده	کل	۸۲۲۴	۱۸۲۴	۶۴۰۰
	جرحی/ فوتی	۳۵۱۱	۲۰۱۰	۱۵۹۱
	خسارتی	۴۷۱۳	۶۲۲	۳۹۹۱

جدول ۷. ماتریس آشفتگی مدل پیشنهادی

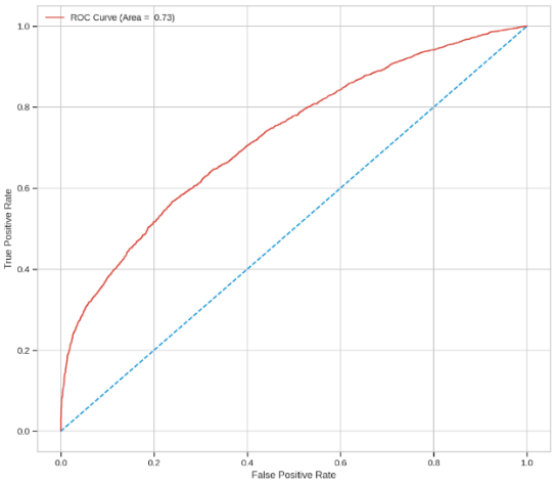
		برآورد		
			جرحی/ فوتی	خسارتی
مشاهده	کل	۸۲۲۴	۱۸۲۴	۶۴۰۰
	جرحی/ فوتی	۳۵۱۱	۲۰۳۰	۱۵۸۱
	خسارتی	۴۷۱۳	۶۲۴	۳۹۸۹

همانطور که از اطلاعات این جدول ها و جدول شماره ۴ مشخص است، عملکرد مدل پیشنهادی در پیش بینی صحیح تصادفات جرحی/فوتی که دارای مشاهدات کمتری نسبت به تصادفات خسارتی است، به نسبت سایر مدل هه بهتر است. مدل پیشنهادی از مجموع ۳۵۱۱ داده جرحی/فوتی، تعداد ۲۰۳۰ تصادف را به درستی تشخیص داده است. این عدد برای مدل شبکه عصبی پیچشی با حذف تصادفی برابر با ۲۰۱۰ و برای مدل لوجیت ترکیبی برابر با ۹۶۳ است.

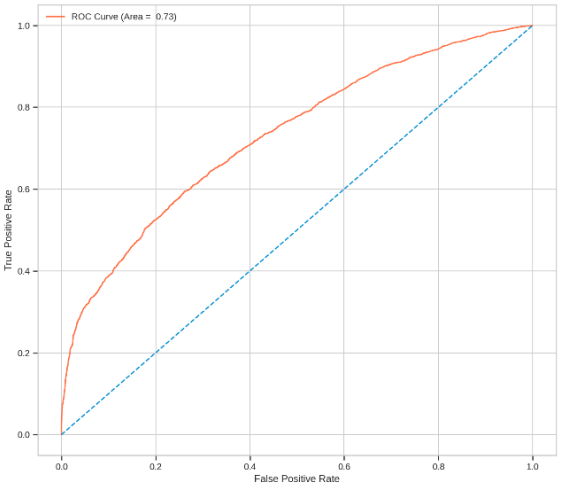
همچنین در ادامه منحنی ROC برای طبقه بندی برای مدل ها ارایه شده است.



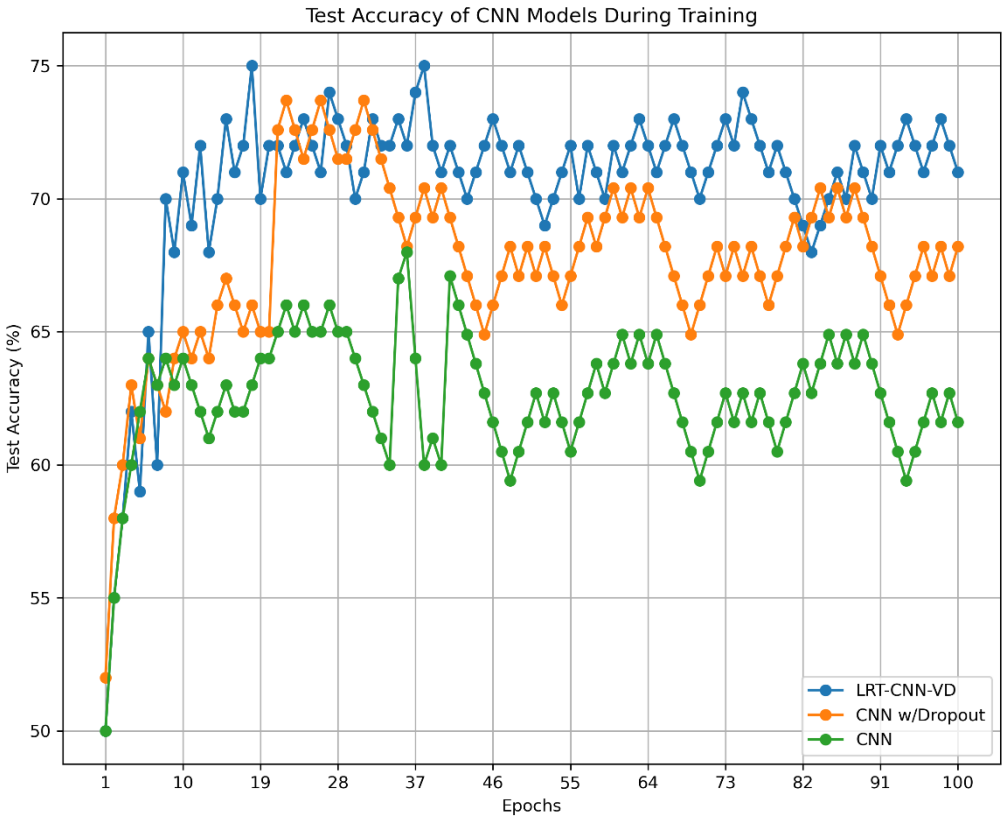
شکل ۴. منحنی ROC مدل لوجیت ترکیبی



شکل ۵. منحنی ROC شبکه عصبی پیچشی با حذف تصادفی



شکل ۶. منحنی ROC مدل پیشنهادی



شکل ۷. نمونه ای از یک دور کامل آموزش مدل های مبتنی بر شکه عصبی پیچشی

با توجه به منحنی ROC در مقایسه بین مدل پیشنهادی و شبکه عصبی پیچشی با مدل لوجیت ترکیبی، دو مساله مورد نظر است. یک اینه مقادیر AUC در هر دو مدل مبتنی بر شبکه عصبی پیچشی از مدل لوجیت ترکیبی بیشتر است. و دوم اینکه هر دو مدل مبتنی بر شبکه عصبی پیچشی، نرخ منفی کاذب کمتری دارند، به این معنا که احتمال دسته بندی اشتباه تصادفات جرحی/فوتی در این مدل ها کمتر است که با توجه به اهمیت بالاتر این سطح از شدت تصادفات، این امر می‌تواند ارزشمند باشد.همچنین تصویر ۷ نشان دهنده یکی از دفعات آموزش کامل شبکه های عصبی پیچشی است. در هر بار آموزش ۱۰۰ دور کامل طی شده است. تصویر نشان دهنده اثر بیش برازش بر مدل بدون نظم دهی است.

۵. نتیجه گیری و پیشنهادات

پژوهش پیش رو سعی در ارایه مدلی مبتنی بر شبکه عصبی پیچشی بیزی داشته است تا با استفاده از روش های تغییر متغیر محلی و نظم دهی حذف تصادفی تغییراتی، برخی از مشکلات موجود در ادبیات موضوع را برطرف کند. ابتدا برای برآورد گرادیان در پس انتشار در مرحله آموزش شبکه عصبی پیچشی بیزی، از روش تغییر متغیر محلی استفاده شد. سپس با استفاده از روش نظم دهی حذف تصادفی تغییراتی، روشی برای یافتن مقدار بهینه ابر پارامتر نظم دهی ارایه شد که این امر از دو جهت مورد توجه است:

۱. اول آنکه مرحله تعیین ابر پارامتر نظم دهی با استفاده از اعتبار سنجی متقابل و جستجو شبکه ای را از بین می‌برد و مقدار بهینه پارامتر نظم دهی در حین مرحله آموزش همراه سایر پارامتر های مدل بدست می‌آید. با توجه به زمان محاسبات بالا شبکه های عصبی، استفاده از این روش می‌تواند فرآیند آموزش را به صورت قابل توجهی تسهیل کند.

۲. دوم آنکه مقدار تعیین شده توسط روش جستجوی شبکه ای برای ابرپارامتر نظم دهی لزوما مقدار بهینه آن نیست. چرا که جستجو در فضای گسسته انجام می‌شود. اما روش پیشنهادی مقدار بهینه این ابرپارامتر را در حین فرآیند آموزش و در فضای پیوسته بدست می‌آورد و در نتیجه عملکرد کلی مدل در مواجهه با بیش برازش بهتر خواهد شد.

در ادامه داده های تصادفات جاده های دو خطه دوطرفه جدا نشده در ایران پس از پیش پردازش بر روی چهار مدل آزمایش شدند. این مدل ها به ترتیب مقابل هستند، مدل لوجیت ترکیب، مدل مبتنی بر شبکه عصبی پیچشی بدون نظم دهی، مدل مبتنی بر شبکه عصبی پیچشی با نظم دهی حذف تصادفی و مدل پیشنهادی. نتیاج نشان داد که عملکرد مدل پیشنهادی چه از حیث دقت و چه از حیث سایر شاخص های طبقه بندی که در این پژوهشی Precision، Recall و F1 Score انتخاب شدند، به طور معنا داری بیشتر بوده است.

بررسی ها نشان دادند که که اگرچه عملکرد مدل شبکه عصبی پیچشی بدون نظم دهی بر روی داده های آموزش بیشتر بوده است، اما به علت بیش برازش عملکرد این مدل به نسبت سایر مدل های مبتنی بر شبکه عصبی پیچشی که از نظم دهی استفاده کرده‌اند بر روی داده های تست به صورت قابل توجهی بیشتر است.

در رابطه با مقایسه دو روش مختلف نظم دهی که استفاده شده نیز شایان ذکر است که مدل پیشنهادی عملکرد بهتری دارد. علت این امر در بهینه سازی ابرپارمتر نظم دهی در روش پیشنهادی است. از آنجا که جستجو برای مقدار بهینه ابرپارامتر نظم دهی در روش حذف تصادفی در فضای گسسته انجام می‌شود، مقدار بدست آمده لزوما مقدار بهینه این ابرپارامتر نیست. اما مدل پیشنهادی مقدار بهینه ابرپارامتر نظم دهی را در حین آموزش پیدا کرده و در نتیجه به نسبت نظم دهی حذف تصادفی عملکرد بهتری داشته است.

در خصوص پیشنهادات برای مطالعات آتی موارد زیر قابل ذکر است:

۱. تحلیل حساسیت متغیر ها بر روی مدل پیشنهادی: اگرچه که شبکه های عصبی بر خلاف مدل های آماری امکان تحلیل پارامتر های مدل را در اختیار نمی‌گذارند، اما روش های مختلفی برای تحلیل حساسیت برای این مدل ها ذکر شده که می تواند در رابطه با مدل پیشنهادی مورد استفاده قرار گیرد، برخی از این روش ها می‌توان به تحلیل SHAP، تحلیل Saliency Maps و تحلیل های مبتنی بر Taylor Decomposition اشاره کرد.

۲. شبکه های عصبی بیزی امکان ارایه معیار های عدم قطعیت در رابطه با برآورد مدل ارایه می‌کنند. این معیار ها امکان بررسی اطمینان مدل از برآورد ها را ارایه می‌کند که در موارد مختلفی مانند تصمیم گیری ها می‌تواند مورد استفاده قرار گیرد. همچنین امکان بررسی منبع عدم قطعیت نیز موجود است، به این منظور که هر یک از عوامل داده و مدل عدم قطعیت موجود در رابطه با یک برآورد چه تاثیری داشته‌اند.

۳. استفاده از روش های یادگیری گروهی برای بهبود عملکرد مدل پیشنهادی. این روش ها با آموزش چند شبکه عملکرد مدل را بهبود می‌بخشند.

۴. در نظر گرفتن اثرات زمانی در مدل سازی با استفاده از شبکه های بازگشتی به همراه شبکه های پیچشی می‌تواند در بهبود عملکرد مدل موثر باشد.

۵. همانطور که اشاره شده، برای جلو گیری از تاثیرات ناشی از تفاوت زیاد مشاهدات در کلاس ها، مشاهدات جرحی و فوتی با هم جمع شده و یک کلاس در نظر گرفته شده اند. انواع روش ها می توانند تاثیر نامتعادل بودن تعداد مشاعدات در کلاس ها را تعدیل کنند، برخی از این روش ها، مبتنی بر الگوریتم های مولد هستند که می‌توانند مورد استفاده قرار گیرند.

۶. مراجع

– Abdelwahab, H. T. and M. A. Abdel-Aty (2001). "Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections." Transportation research record 1746(1): 6-13.

– Ali, A., Y. Zhu and M. Zakarya (2022). "Exploiting dynamic spatio-temporal graph convolutional neural networks for citywide traffic flows prediction." Neural networks 145: 233-247.

– Bejani, M. M. and M. Ghatee (2019). "Regularized deep networks in intelligent transportation systems: A taxonomy and a case study." arXiv preprint arXiv:1911.03010.

– Bejani, M. M. and M. Ghatee (2021). "A systematic review on overfitting control in shallow and deep neural networks." Artificial Intelligence Review: 1-48.

– Blei, D. M., A. Kucukelbir and J. D . McAuliffe (2017). "Variational inference: A review for statisticians." Journal of the American statistical Association 112(518): 859-877.

– Chimba, D. and T. Sando (2009). "The prediction of highway traffic accident injury severity with neuromorphic techniques." Advances in transportation studies 19: 17-26.

– Delen, D., R. Sharda and M. Bessonov (2006). "Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks." Accident Analysis & Prevention 38(3.444-434).

– Do, L. N., H. L. Vu, B. Q. Vo, Z. Liu and D. Phung (2019). "An effective spatial-temporal attention based neural network for traffic flow prediction." Transportation research part C: emerging technologies 108: 12-28.

– Fang, S.-H., Y.-X. Fei, Z .Xu and Y. Tsao (2017).

– "Learning transportation modes from smartphone sensors based on deep neural network." IEEE Sensors Journal 17(18): 6111-6118.

– Fanyu, M., N. Sze, S. Cancan, C. Tiantian and Z. Yiping (2021). "Temporal instability of truck volume composition on non-truck-involved crash severity using uncorrelated and correlated grouped random parameters binary logit models with space-time variations." Analytic Methods in Accident Research 31: 100168.

– Haghshenas, S. S., G. Guido, A. Vitale and V. Astarita (2023). "Assessment of the level of road crash severity: Comparison of intelligence studies." Expert Systems with Applications 234: 121118.

– Hou, Q., X. Huo, J. Leng and F. Mannering (2022). "A note on out-of-sample prediction, marginal effects computations, and temporal testing with random parameters crash-injury severity models." Analytic methods in accident research 33: 100191.

– Hu, Z., J. Zhou, K. Huang and E. Zhang (2022). "A data-driven approach for traffic crash prediction: A case study in Ningbo, China." International Journal of Intelligent Transportation Systems Research 20(2): 508-518.

– Islam, M. R., K. B. Kelarestaghi, A. Ermagun and S. Banerjee (2019). "Gender Differences in Injury Severity Risk of Single-Vehicle Crashes in Virginia: A Nested Logit Analysis of Heterogeneity." arXiv preprint arXiv:1901.03289.

recurrent neural networks." Applied Sciences 7(6): 476.

– Shiran, G., R. Imaninasab and R. Khayamim (2021). "Crash severity analysis of highways based on multinomial logistic regression model ,decision tree techniques, and artificial neural network: a Modeling comparison." Sustainability 13(10): 5670.

– Silva, P. B., M. Andrade and S. Ferreira (2020). "Machine learning applied to road safety modeling: A systematic literature review." Journal of traffic and transportation engineering (English edition) 7(6): 775-790.

– Vajari, M. A., K. Aghabayk, M. Sadeghian and N. Shiwakoti (2020). "A multinomial logit model of motorcycle crash severity at Australian intersections." Journal of safety research 73: 124-7.

– Yu, H., Z. Li, G. Zhang, P. Liu and T. Ma (2021). "Fusion convolutional neural network-based interpretation of unobserved heterogeneous factors in driver injury severity outcomes in single-vehicle crashes." Analytic methods in accident research 30 :100157.

– Zeng, Q., W. Gu, X. Zhang, H. Wen, J. Lee and W. Hao (2019). "Analyzing freeway crash severity using a Bayesian spatial generalized ordered logit model with conditional autoregressive priors." Accident Analysis & Prevention 127: 87-95.

– Zeng, Q. and H. Huang (2014). "A stable and optimized neural network model for crash injury severity prediction." Accident Analysis & Prevention 73: 351-358.

– Kingma, D. P., T. Salimans and M. Welling (2015). "Variational dropout and the local reparameterization trick." Advances in neural information processing systems 28.

– Li, L., J. Zhu, H. Zhang, H. Tan, B. Du and B. Ran (2020). "Coupled application of generative adversarial networks and conventional neural networks for travel mode detection using GPS data." Transportation Research Part A: Policy and Practice 136: 282-292.

– Li, P., M. Abdel-Aty and J. Yuan (2020). "Real-time crash risk prediction on arterials based on LSTM-CNN." Accident Analysis & Prevention 135: 105371.

– Park, S. and J. Park (2022). "Multilevel Mixed-Effects Models to Identify Contributing Factorson Freight Vehicle Crash Severity." Sustainability 14(19): 11804.

– Patil, S., S. R. Geedipally and D. Lord (2012). "Analysis of crash severities using nested logit model—accounting for the underreporting of crashes." Accident Analysis & Prevention 45: 646-653.

– Rahim, M. A. and H. M .Hassan (2021). "A deep learning based traffic crash severity prediction framework." Accident Analysis & Prevention 154: 106090.

– Rezapour, M., S. Nazneen and K. Ksaibati (2020). "Application of deep learning techniques in predicting motorcycle crash severity." Engineering Reports 2(7): e12175.

– Rezapour, M., S. S. Wulff and K. Ksaibati (2020). "Bayesian hierarchical modelling of traffic barrier crash severity." International journal of injury control and safety promotion 28(1): 94-102.

– Sameen, M. I. and B .Pradhan (2017). "Severity prediction of traffic accidents with