

طراحی و پیاده‌سازی سیستم هوش مصنوعی توضیح‌پذیر برای خودروهای هوشمند

چکیده

با پیشرفت‌های قابل توجه در زمینه‌ی یادگیری عمیق برای سیستم‌های خودران، چالش‌های مهمی در زمینه شفافیت، قابلیت تفسیر و پذیرش کاربران همچنان باقی مانده است. عدم توانایی در توضیح رفتار تصمیم‌گیری مدل‌های یادگیری عمیق، امکان پیاده‌سازی عملی آن‌ها را در سناریوهای مختلف دنیای واقعی محدود می‌کند که این موضوع در شبکه‌های تصمیم‌گیری حیاتی است؛ زیرا نتایج نادرست ممکن است منجر به حوادث خطرناک جاده‌ای شود. این پایان‌نامه رویکردی نوآورانه برای ایجاد سیستم‌های خودران قابل تفسیر ارائه می‌دهد که شامل سه نوآوری اصلی است. ابتدا، یک مدل معرفی می‌شود که ماژول توجه بلوک پیچشی را با توجه عمیق متصل ترکیب می‌کند و با یک مدل تصمیم‌گیرنده و توضیح‌دهنده‌ی متنی پیشرفته، یکپارچه شده است. این ترکیب، منجر به ایجاد سیستمی دقیق‌تر و جامع‌تر برای تفسیرپذیری می‌شود. دومین نوآوری این تحقیق، ارائه رویکرد یادگیری چندوظیفه برای پیش‌بینی همزمان رفتارهای مختلف رانندگی و دلایل زیربنایی آن‌ها است. هم‌چنین، تحلیل حساسیت آستانه اطمینان برای مدل‌های طبقه‌بندی چندوظیفه انجام شده که نشان می‌دهد رویکردهای آستانه‌ی ثابت سنتی برای سناریوهای چندوظیفه بهینه نیستند؛ علاوه‌براین، مجموعه داده IUST-XAI-AD شامل ۹۵۸ تصویر با حاشیه‌نویسی‌های انسانی تفصیلی برای اقدامات رانندگی و استدلال تبیینی معرفی می‌شود. این مجموعه داده، شکاف مهمی در معیارهای ارزیابی خاص حوزه برای بافت‌های رانندگی فارسی پر می‌کند و محیط آزمایش چالش‌برانگیزتری نسبت به مجموعه داده‌های موجود فراهم می‌کند. ارزیابی‌های کمی با استفاده از معیار F1 استاندارد بر روی مجموعه داده BDD-OIA، nu-AR و مجموعه داده جدید IUST-XAI-AD انجام شده است. نتایج تجربی نشان می‌دهد که تکنیک پیشنهادی با دقت ۵۴.۰۶٪ و ۳۶.۶۵٪ به ترتیب در overall F1 score و mean F1 score روی مجموعه داده‌ی BDD-OIA از مدل‌های موجود بهتر عمل کرده و بهبود قابل توجهی در قابلیت تفسیر نشان می‌دهد. تحلیل حساسیت آستانه اطمینان به طور قابل توجهی عملکرد مدل را افزایش می‌دهد، در حالی که معرفی مجموعه داده IUST-XAI-AD بینش‌های مهمی درباره الگوهای رفتار رانندگی میان فرهنگی آشکار می‌کند و استحکام و تعمیم‌پذیری مدل ارائه‌شده را نسبت به مدل‌های موجود نشان می‌دهد. این تحقیق با ارائه پیشرفت‌های روش‌شناختی و ابزارهای ارزیابی عملی، به توسعه سیستم‌های خودران قابل اعتمادتر و قابل تفسیرتر کمک می‌کند و زمینه‌ای برای آینده‌ی وسایل نقلیه خودران قابل تفسیر فراهم می‌آورد.

واژه‌های کلیدی: یادگیری چندوظیفه، تحلیل حساسیت آستانه اطمینان، وسایل نقلیه خودران، هوش مصنوعی قابل تفسیر، بینایی ماشین، اعتبارسنجی مجموعه داده، شبکه‌های توجه، توضیح زبان طبیعی

دانشجو: مریم سادات حسینی آزاد

استاد راهنما: دکتر شهریار برادران شکوهی

هیات داور: دکتر ستار میرزا کوچکی؛ دکتر احمد عقیفی

ساعت: ۰۹:۰۰

تاریخ دفاع: چهارشنبه ۱۴۰۴/۷/۳۰

محل: سالن خوارزمی دانشکده برق