

روش یادگیری فدرال غیرمتمرکز مبتنی بر گراف پویا و تصمیم‌گیری چندمعیاره برای شبکه‌های متحرک

نیما کریمی^۱، محمدرضا پژوهان^{۲*}، امیر جهانگرد رفسنجانی^۳

^۱ دانشجوی دکتری، دانشکده مهندسی کامپیوتر، دانشگاه یزد، یزد، ایران

^۲ دانشیار، دانشکده مهندسی کامپیوتر، دانشگاه یزد، یزد، ایران

^۳ استادیار، دانشکده مهندسی کامپیوتر، دانشگاه یزد، یزد، ایران

مقاله پژوهشی

چکیده

تاریخ دریافت:

۱۴۰۴/۱۰/۰۸

تاریخ پذیرش:

۱۴۰۴/۱۲/۲۶

کلیدواژه‌ها:

یادگیری فدرال غیرمتمرکز،
شبکه‌های همتابه‌همتا متحرک،
مجموعه برتری گراف، روش
تصمیم‌گیری چندمعیاره

نویسنده مسئول:

pajoohan@yazd.ac.ir

یادگیری فدرال به عنوان یک رویکرد نوین در یادگیری ماشین توزیع شده، امکان آموزش مدل‌های هوش مصنوعی را بدون نیاز به انتقال داده‌های خام فراهم می‌سازد. با این حال، ساختار کلاسیک یادگیری فدرال وابسته به یک سرور مرکزی است که در بسیاری از کاربردهای واقعی نظیر شبکه‌های متحرک و اینترنت اشیا در دسترس نیست. از سوی دیگر، ناهمگونی داده‌ها و تحرک گره‌ها چالش‌های اساسی در پیاده‌سازی چنین دستگاه‌هایی ایجاد می‌کند. در این پژوهش، یک روش نوین برای یادگیری فدرال غیرمتمرکز ارائه شده است. در این روش، هر گره تنها بر اساس ارتباطات محلی خود و بدون وابستگی به گره مرکزی اقدام به به‌روزرسانی مدل می‌کند. یک گراف پویا مبتنی بر ملاقات گره‌ها برای مدل‌سازی ارتباطات ایجاد شده و به کمک محاسبه مجموعه برتری گراف، گره‌های مؤثر در فرآیند تجمیع انتخاب می‌شوند. برای وزندهی دقیق به مدل‌های منتخب، از یک روش تصمیم‌گیری چندمعیاره استفاده شده است که معیارهایی نظیر دقت، حجم داده و امتیاز گره را در نظر می‌گیرد. علاوه بر این، از شباهت بین مدل‌ها در فرآیند تجمیع گراف‌ها استفاده می‌شود. روش پیشنهادی بر روی سه دادگان مرجع MNIST، FASHION-MNIST و CIFAR10 ارزیابی و با دو روش مرجع FedAvg همتابه‌همتا و WAFL مقایسه شد. نتایج دقت مدل‌ها در دو محیط با ۵۰ و ۱۰۰ گره به ترتیب ۰/۹۶۴، ۰/۹۵۲، ۰/۸۷۱، ۰/۸۸۲، ۰/۴۲۴ و ۰/۳۹۹ نشان داد که روش پیشنهادی عملکرد بهتری دارد. یافته‌ها بیانگر توانمندی روش پیشنهادی برای کاربرد در شبکه‌های پویا و متحرک با تعداد مشتری‌های زیاد و فاقد زیرساخت مرکزی است که می‌تواند بر مشکل مقیاس‌پذیری و انحراف مشتری‌ها در یادگیری فدرال غیرمتمرکز غلبه کند.

doi : 10.22034/ABMIR.2026.24181.1213

۱- مقدمه

انتخاب تصادفی همسایگان برای تجمیع معمولاً منجر به مشارکت گره‌های کم‌کیفیت یا دارای انحراف بالا می‌گردد. برای رفع این چالش‌ها، رویکردهای متعددی از جمله نحوه انتخاب مشتری، وزن‌دهی پویا، خوشه‌بندی محلی و استفاده از معیارهای شباهت مدل‌های یادگیری عمیق پیشنهاد شده است [۶]. با این حال، اکثر این روش‌ها یا همچنان به یک هماهنگ‌کننده مرکزی وابسته‌اند یا در شبکه‌های کاملاً غیرمتمرکز و پویا و بدون هیچ توپولوژی مشخص، کارایی خود را از دست می‌دهند. در بیشتر مطالعات انجام شده در زمینه یادگیری فدرال غیرمتمرکز، شبکه هم‌تابه‌متا دارای یک توپولوژی مشخص در نظر گرفته شده است، لذا مطالعه در محیط کاملاً پویا که مشتریان آزادانه حرکت می‌کنند و توپولوژی مشخصی وجود ندارد، برای رسیدن به رویکردهای عملی و واقعی حائز اهمیت است.

در این مقاله، یک الگوریتم دومرحله‌ای نوین مبتنی بر مفهوم «مجموعه برتری وزن‌دار گراف» ارائه می‌شود که به صورت کاملاً محلی و بدون نیاز به سرور مرکزی، زیرمجموعه‌ای بهینه از مشتری‌ها را برای مشارکت در هر دور تجمیع انتخاب می‌کند. این انتخاب بر مبنای معیارهای ترکیبی کیفیت مدل محلی (دقت)، حجم و شباهت مدل‌های عمیق با استفاده از شباهت کسینوسی و همبستگی پیرسون انجام می‌گیرد. برای وزن‌دهی به گره‌های منتخب، از رویکرد تصمیم‌گیری چندمعیاره استفاده شده است. سپس تجمیع وزن‌دار تنها روی این گره‌های صورت می‌پذیرد.

ارزیابی نتایج بر روی داده‌های MNIST Fashion-، CIFAR10 برای ۵۰ و ۱۰۰ مشتری در شبکه پویای متحرک نشان می‌دهد که روش پیشنهادی در مقایسه با دو روش دیگر به‌طور متوسط بهبود ۳/۷ و ۶/۱ درصدی در دقت مدل ایجاد می‌کند. این نتایج بیانگر توانایی بالای الگوریتم پیشنهادی در حل اساسی مشکلات مقیاس‌پذیری و انحراف مشتری^۸ در یادگیری فدرال کاملاً غیرمتمرکز است. ساختار ادامه این مقاله به این صورت

یادگیری فدرال^۱ به‌عنوان یک رویکرد نوین یادگیری ماشینی، امکان آموزش مدل‌ها را بدون تبادل مستقیم داده‌های خام فراهم می‌آورد و بدین ترتیب حریم خصوصی کاربران را حفظ می‌کند [۱]. با گسترش کاربردهای هوش مصنوعی در دستگاه‌های لبه، اینترنت اشیاء و شبکه‌های حسگر، نیاز به نسخه‌های کاملاً غیرمتمرکز این رویکرد که فاقد یک سرور اصلی باشند، روزبه‌روز افزایش یافته است؛ این افزایش استفاده از یادگیری فدرال غیرمتمرکز تا حد زیادی ناشی از گسترش دستگاه‌های اینترنت اشیاء و شبکه‌های ارتباط بین‌خودرویی در سال‌های اخیر است؛ زیرا این سامانه‌ها داده‌های حساس و توزیع شده تولید می‌کنند که انتقال آن‌ها به یک سرور مشخص به خاطر حریم خصوصی، پهنای باند و تأخیر کارآمد نیست [۲]. در یادگیری فدرال غیرمتمرکز^۲ مبتنی بر شبکه‌های هم‌تابه‌متا، گره‌ها صرفاً از طریق ارتباطات محلی و پروتکل‌های مبتنی بر شایعه یا مشابه آن با یکدیگر هماهنگ می‌شوند [۳]. این رویکرد از طرفی با حذف نقطه تکین خرابی^۳، که در واقع بخشی از یک سیستم است که با ازکارافتادن آن کل سیستم از کار می‌افتد، مقیاس‌پذیری بسیار بالایی ارائه می‌دهد، اما از طرف دیگر با چالش‌های جدی در زمینه همگرایی، به‌ویژه در حضور ناهمگونی داده‌ها و پویایی توپولوژی شبکه مواجه است [۳].

روش کلاسیک FedAvg که ابتدا توسط مک‌مهن و همکاران معرفی شد [۱]، در شرایط توزیع غیرمستقل و غیرهم‌توزیع داده‌ها^۴ و نیز افزایش تعداد مشتری‌ها، دچار افت شدید دقت و همگرایی ضعیف می‌شود [۴]. مطالعات متعدد نشان داده‌اند که در سناریوهای با بیش از ۵۰ مشتری و مقادیر کوچک پارامتر توزیع دیریکله^۵، دقت مدل در FedAvg غیرمتمرکز می‌تواند تا ۳۰ درصد یا بیشتر کاهش یابد [۵]. این مشکل در شبکه‌های هم‌تابه‌متای پویا که توپولوژی به‌صورت مداوم تغییر می‌کند (مانند شبکه‌های موردی متحرک^۶ با مدل متحرک نقطه تصادفی مقصد^۷) تشدید می‌شود، زیرا

⁵ Dirichlet Distribution

⁶ Mobile Adhoc Network

⁷ Random Waypoint

⁸ Client Drift

¹ Federated Learning

² Decentralized Federated Learning

³ Single Point of Failure

⁴ Non-Independent and Identically Distributed(Non-IID)

اثرات منفی این نوع داده‌ها را حتی در شبکه‌های موردی متحرک کاهش می‌دهد. این الگوریتم روی داده‌های واقعی غیرمستقل و غیرهم‌توزیع تست شده است و نتایج نشان‌دهنده بهبود دقت و نیرومندی در مقایسه با روش‌های مرکزی یا غیرمتمرکز است.

کین و همکاران چارچوب BlockDFL را بر پایه بلاکچین و پروتکل شایعه^۱ معرفی کرده‌اند که با استفاده از اثبات سهام، تجمع مدل‌ها را بدون سرور مرکزی مدیریت می‌کند [۹]. ارزیابی روی CIFAR10 و ImageNet با مدل ResNet-18 به دقت ۸۷٪ و کاهش ۳۰٪ هزینه ارتباطی نسبت به FedAvg منجر شد. مزایای این روش شامل امنیت بالا با ثبت غیرقابل تغییر، مقیاس‌پذیری برای هزاران گره و حفظ حریم خصوصی از طریق رمزنگاری است، اما با چالش‌هایی مانند تأخیر تجمع، مصرف انرژی بلاکچین و پیچیدگی پیاده‌سازی مواجه است.

چارچوب SCAMA^۲ برای یادگیری فدرال غیرمتمرکز در محیط‌های همتابه‌متا طراحی شده و بر همگام‌سازی و تجمع مدل‌ها در کاربردهایی مانند طبقه‌بندی تصاویر پزشکی با داده‌های غیرمستقل و غیرهم‌توزیع تمرکز دارد [۱۰]. این روش با دقت ۹۲٪ بهبود ۱۲٪ نسبت به FedAvg و همگرایی سریع‌تر عملکرد مؤثری در سناریوهای با تأخیر ارتباطی و ناهمگونی داده نشان داده است. چن و همکاران یک چارچوب یادگیری فدرال را معرفی می‌کنند که این رویکرد نوین برای محیط‌های غیرمتمرکز و بر پایه وضعیت بلاکچین ساخته شده است [۱۱]. کاربرد اصلی این مدل در طبقه‌بندی تصاویرمانند تشخیص الگوهای تصویری در شبکه‌های لبه است که داده‌های حساس و غیرمستقل و غیرهم‌توزیع را در برمی‌گیرد. این مدل دقت طبقه‌بندی تصاویر تا ۹۴٪ روی MNIST و ۸۵٪ روی CIFAR100، با همگرایی در ۴۰-۶۰ دور و بهبود ۲۰٪ نسبت به FedAvg استاندارد داشت. محققان یک بررسی جامع و تجربی از چالش‌های ناهمگونی در یادگیری فدرال با تمرکز بر محیط‌های همتابه‌متا و کاربردهای محاسباتی هوش مصنوعی در شبکه‌های بی‌سیم ارائه داده‌اند [۱۲]. آن‌ها بر انواع ناهمگونی مانند داده‌های غیرمستقل و غیرهم‌توزیع، تفاوت‌های سخت‌افزاری مشتری‌ها و مدل‌ها تأکید داشتند. الگوریتم پیشنهادی آن‌ها یک

است که در بخش دوم به بررسی روش‌های ارائه‌شده قبلی پرداخته می‌شود، سپس در بخش سوم روش پیشنهادی به تفصیل توضیح داده خواهد شد. در بخش بعدی نتایج به‌دست‌آمده ارائه و تفسیر می‌شود و در بخش آخر هم نتیجه‌گیری و پیشنهادها برای گام‌های بعدی بیان می‌شود.

۲- پیشینه تحقیق و بیان چالش‌ها

استفاده رو به رشد یادگیری فدرال و مزایای آن، در سال‌های اخیر منجر به انجام مطالعات گسترده‌ای در این حوزه شده است. یادگیری فدرال استاندارد، داده‌های خصوصی کاربران را حفظ می‌کند زیرا داده‌ها روی دستگاه‌های محلی مانند گوشی‌ها یا دستگاه‌های اینترنت اشیا آموزش می‌بینند و فقط به‌روزرسانی‌های مدل به یک سرور ارسال می‌شود. با این حال، این مدل مرکزی با مشکلاتی نظیر وابستگی به سرور قدرتمند و محدودیت‌های مقیاس‌پذیری مواجه است که محققان را به سمت راهکارهای غیرمتمرکز سوق داده است.

کاراس و همکاران [۷] یک رویکرد یادگیری فدرال کاملاً غیرمتمرکز برای شبکه همتابه‌متا معرفی کرده‌اند که در آن دستگاه‌های لبه بدون نیاز به سرور، مستقیماً مدل‌ها را مبادله و میانگین‌گیری می‌کنند. این روش با حفظ حریم خصوصی، کاهش هزینه‌های ارتباطی، افزایش مقیاس‌پذیری و مقاومت در برابر شکست، برای محیط‌های اینترنت اشیا و کلان‌داده ایده‌آل است. نویسندگان دو الگوریتم نوین پیشنهاد کرده‌اند که روی داده‌های واقعی آزمایش شده و نتایج نشان‌دهنده مقیاس‌پذیری و نیرومندی خوب آن‌ها است.

محققان یک الگوریتم یادگیری فدرال همتابه‌متا در شبکه‌های موردی متحرک پیشنهاد داده‌اند که دستگاه‌ها مستقیماً با همسایگان خود ارتباط برقرار می‌کنند، مدل‌ها را به اشتراک می‌گذارند و تجمع می‌کنند [۸]. چالش اصلی در این روش داده‌های غیرمستقل و غیرهم‌توزیع است. در واقعیت، داده‌های محلی دستگاه‌ها اغلب غیرهمگون هستند، که باعث کاهش دقت مدل سراسری می‌شود. این پژوهش چند راهبرد جدید تجمع مدل پیشنهاد می‌کند که

² Smart-Contract-Driven Asynchronous Model Aggregation

¹ Gossip protocol

وزن $FEWS^1$ و $OEWS^2$ تمرکز دارد. این مطالعه نشان داد که $OEWS$ دقت طبقه‌بندی تصاویر را تا ۸۸٪-۹۲٪ افزایش می‌دهد. ارزیابی روی دادگان‌های خاص و سفارشی از کارخانه‌های تولیدی، نشان‌دهنده نیرومندی بیشتر روش جدید در مقابل داده‌های پرت نسبت به روش وزن‌دهی یکنواخت است.

دوتا و همکاران [۱۶] یک رویکرد توزیع‌شده مبتنی بر یادگیری فدرال برای تشخیص و مکان‌یابی جنگل‌زدایی با تصاویر ماهواره‌ای Landsat ارائه داده‌اند که با آموزش مشارکتی بین سازمان‌های منطقه‌ای، مشکلات حریم خصوصی و انتقال داده‌های حجیم را حل کرده و به دقت ۹۰٪ دست یافته است. همچنین چارچوب یادگیری فدرال همتابه‌ممتا چندوجهی که تصاویر با فرمت‌های مختلف را برای طبقه‌بندی مناظر و پهنه‌های زمین ادغام می‌کند توسط محققان معرفی شده است [۱۷]. این روش شامل انکودرهای مخصوص حالت^۳ در شبکه همتابه‌ممتا است. طبق نتایج ارزیابی روی داده‌های EuroSAT، روش پیشنهادی به دقت ۹۲٪ و بهبود ۱۰٪ دست یافته است. علاوه بر این یک روش یادگیری فدرال همتابه‌ممتای آنلاین برای طبقه‌بندی چندوظیفه‌ای تصاویر در مطالعه‌ای انجام شده است [۱۸]. در این مطالعه یک روش که شامل امتیاز قابلیت اعتماد با استنتاج بیزی است. مدل روی KITTI dataset به دقت ۸۹٪ دست یافته است.

علیرغم مزایای یادگیری فدرال در حفظ حریم خصوصی داده‌ها، روش‌های موجود با چالش اساسی مواجه هستند. معماری متمرکز به دلیل وابستگی به یک سرور اصلی با مشکلاتی نظیر نقطه تکین خرابی و محدودیت مقیاس‌پذیری در سناریوهای واقعی مواجه است. در مقابل، رویکردهای غیرمتمرکز همتابه‌ممتا که برای رفع این محدودیت‌ها معرفی شده‌اند، با چالش ناهمگونی داده‌ها در دستگاه‌های مختلف مواجه است که باعث کاهش دقت مدل سراسری و مشکلات همگرایی می‌شود. از طرفی وجود توپولوژی ثابت در یک شبکه متحرک از جنبه عملی، امری غیرواقعی است که بیشتر مطالعات انجام‌شده پیشین با این فرض انجام شده‌اند. علاوه بر این، تجمع مدل‌ها بدون توجه به سابقه ملاقات مشتریان و کارایی مدل آن‌ها می‌تواند ارزش مدل‌ها را در فرآیند تجمع تحت تأثیر

روش تطبیق همسایگی همتابه‌ممتاست، که مشتری‌ها را بر اساس شباهت داده‌ها با معیارهایی مانند فاصله کسینوسی گرادیان یا واگرایی کولبک-لیبر جفت می‌کند. این روش شامل خوشه‌بندی پویا با روش K-Means است. در طبقه‌بندی تصاویر FEMNIST، دقت مدل سراسری تا ۸۹٪ رسید که بهبود ۱۵٪ نسبت به روش FedAvg در داده‌های ناهمگون داشت.

چالش‌های ارتباطی در یادگیری فدرال غیرمتمرکز برای داده‌های گراف به‌ویژه در سناریوهای همتابه‌ممتا با داده‌های غیرمستقل و غیرهم‌توزیع در مورد بررسی قرار گرفته است [۱۳]. نویسندگان به کاربرد طبقه‌بندی تصاویر مبتنی بر گراف پرداخته‌اند و از گراف‌ها برای مدل‌سازی روابط مانند تعبیه گره در شبکه‌های اجتماعی با ویژگی‌های تصویری استفاده کرده‌اند. رویکرد مقاله ترکیبی از فشرده‌سازی و تجمع غیرمتمرکز است. این مدل دقت طبقه‌بندی را تا ۹۱٪ روی دادگان CORA بهبود داده است که حدوداً ده درصد بهتر از FedAvg است. همچنین به دقت ۸۷٪ روی دادگان PubMed، با همگرایی در ۴۰ دور در مقابل ۸۰ دور مدل پایه دست یافت.

پژوهشگران چارچوبی برای یادگیری فدرال کاملاً توزیع‌شده و بدون گره مرکزی ارائه می‌دهند که در آن گره‌ها در یک شبکه و به‌صورت فرصت‌طلبانه با یکدیگر تعامل دارند [۱۴]. در این روش، هر گره مدل را با استفاده از داده‌های محلی خود آموزش داده و در زمان برقراری ارتباط با سایر گره‌های مجاور، مدل‌های محلی میان آن‌ها تبادل و تجمع می‌شود. نتایج این پژوهش نشان می‌دهد که این رویکرد در محیط‌های با اتصال ناقص و در حضور داده‌های غیرمستقل و غیرهم‌توزیع می‌تواند از واگرایی مدل‌های محلی جلوگیری کرده و به بهبود تعمیم‌پذیری مدل منجر شود.

یک مطالعه تجربی جامع بر روی راهبردهای انتخاب وزن در یادگیری فدرال با تمرکز بر طبقه‌بندی تصاویر در حوزه تولیدات صنعتی انجام شده است [۱۵]. نویسندگان بر چالش‌های خاص این حوزه مانند حفظ حریم خصوصی داده‌های صنعتی، ناهمگونی داده‌ها و تأثیر انتخاب وزن بر همگرایی مدل سراسری تأکید دارند. رویکرد مقاله کاملاً تجربی است و بر مقایسه دو استراتژی انتخاب

³ Modality-specific encoders

¹ Final Epoch Weight Selection

² Optimal Epoch Weight Selection

۲) پوشش خارجی کامل: هر رأس خارج از مجموعه حداقل با یک رأس داخل مجموعه مجاور است (یعنی مجموعه تمام گره‌های شبکه را پوشش می‌دهد).

مفهوم کلاسیک مجموعه برتری گراف توسط پاون و پیلو به صورت رسمی معرفی و به مسائل خوشه‌بندی، انتخاب زیرگروه و پوشش گره تعمیم داده شده است [۱۹]. در این پژوهش، از نسخه وزن دار و پیاده‌سازی حریصانه مجموعه برتری برای انتخاب بهینه گره‌های مشارکت کننده در فرآیند تجمیع استفاده می‌شود.

۳-۱-۲ شباهت مدل‌های یادگیری عمیق^۳

شباهت‌سنجی بین مدل‌های یادگیری عمیق در سال‌های اخیر به یکی از موضوعات مهم در تحلیل و درک رفتار شبکه‌های عصبی تبدیل شده است. ایده اصلی این حوزه آن است که بررسی کنیم دو مدل تا چه حد از نظر نحوه یادگیری و تصمیم‌گیری به یکدیگر نزدیک هستند، حتی اگر معماری متفاوتی داشته باشند یا به روش‌های متفاوتی آموزش دیده باشند [۲۰]. این نوع تحلیل‌ها کمک می‌کند تا پدیده‌هایی مانند میزان یادگیری، تنوع میان مدل‌ها و همگرایی بهتر فهمیده شود. در مجموع، سنجش شباهت میان مدل‌های یادگیری عمیق ابزار مهمی برای درک عمیق‌تر فرآیند یادگیری و طراحی مدل‌های مؤثرتر به شمار می‌آید و نقش مکملی در کنار معیارهای متداول ارزیابی عملکرد ایفا می‌کند.

۳-۲ الگوریتم روش پیشنهادی

الگوریتم پیشنهادی در هر دور یادگیری فدرال شامل دو مرحله اصلی است، که در مرحله اول عملیات انتخاب گره‌ها انجام می‌شود. در مرحله دوم وزن هر یک از گره‌های انتخاب شده برای میزان مشارکت آن گره در فرآیند تجمیع و میانگین‌گیری محاسبه می‌شود.

۳-۲-۱ انتخاب گره با استفاده از مجموعه برتری

هر گره به صورت محلی یک گراف وزن دار از گره‌های ملاقات شده تا لحظه فعلی را می‌سازد و نگهداری می‌کند. در هر ملاقات، دو گره گراف‌های محلی خود را با یکدیگر به اشتراک گذاشته و گراف

قرار دهد و منجر به کاهش کارایی و همگرایی مدل سراسری شود. در این پژوهش بر اساس دیدارهای پیشین مشتریان و میزان شباهت مدل‌های آن‌ها، یک گراف ارتباطی منحصربه‌فرد برای هر مشتری ساخته می‌شود؛ در ادامه با محاسبه مجموعه برتری گراف، مشتریانی که مدل‌های کارتری دارند برای عملیات تجمیع انتخاب می‌شوند. سپس با استفاده از روش‌های تصمیم‌گیری چندمعیاره وزن مدل‌های برتر برای تجمیع تعیین می‌شود.

۳-۲ روش پیشنهادی

در این بخش روش پیشنهادی که دربرگیرنده مفاهیم و روابط ریاضی، الگوریتم و ترتیب انجام مراحل است، بیان می‌شود. این روش برای استفاده در شبکه همتابه‌متای کاملاً غیرمتمرکز و پویا ارائه شده است که در دو مرحله فرآیند تجمیع را انجام می‌دهد، در مرحله اول با استفاده از ساخت گراف همسایگی و محاسبه مجموعه برتری گراف^۱ گره‌های شرکت کننده در فرآیند تجمیع انتخاب می‌شوند و در مرحله دوم وزن هر یک از گره‌های انتخاب شده در فرآیند تجمیع مدل‌ها محاسبه می‌شود. در ادامه این بخش، ابتدا مفاهیم مورد استفاده و سپس مراحل روش پیشنهادی بیان می‌شوند.

۳-۱ مفاهیم کاربردی

توسعه روش پیشنهادی بر مبنای دو مفهوم اصلی است. اولین مورد مجموعه برتری در تئوری گراف است و دیگری شباهت مدل‌های یادگیری عمیق و شبکه‌های عصبی است که هر دو مورد در ادامه توضیح داده می‌شوند.

۳-۱-۱ مجموعه برتری وزن دار گراف^۲

در نظریه گراف، یک مجموعه برتری وزن دار زیرمجموعه‌ای از رئوس گراف است به طوری که دارای دو شرط زیر باشد:

۱) وزن داخلی بالا: گره‌های داخل مجموعه دارای بالاترین وزن هستند.

³ Model Similarity

¹ Dominating Set

² Weighted Dominating Set

$$d_{\text{pearson}} = 1 - r(M_i, M_j) \quad (4)$$

پس از نرمال کردن فرمول‌های فاصله، رابطه نهایی فاصله دو مدل به وسیله رابطه (۵) محاسبه می‌شود.

$$d_{\text{final}} = \lambda d_{\text{cosine}} + (1 - \lambda) d_{\text{pearson}} \quad (5)$$

که در آن $\lambda \in [0,1]$ یک پارامتر تنظیمی است که میزان مشارکت هر یک از روش‌ها را در فاصله نهایی مشخص می‌کند. در نهایت، امتیاز هر گره^۱ با استفاده از رابطه (۶) محاسبه می‌شود.

$$\text{score}(v) = \begin{cases} w_v \text{ if } N(v) = \emptyset \\ w_v \cdot \left(\min \left\{ 1, \left(\frac{\text{deg}_v}{\max(\text{deg}_{N(v)})} \right) \right\} \right) & \text{otherwise} \\ 1 + \frac{1}{|N(v)|} \sum_{u \in N(v)} d(v, u) \cdot w_u \end{cases} \quad (6)$$

که در آن w_v وزن خود گره، $N(v)$ مجموعه همسایه‌های گره v ، $d(v, u)$ فاصله بین مدل‌های گره (u, v) و w_u وزن گره همسایه u ، deg_v درجه گره v و $\text{deg}_{N(v)}$ مجموعه درجه گره‌های همسایه گره v هستند. وزن‌ها و فاصله‌ها نرمال‌سازی شده‌اند، همچنین در مدل‌های بزرگ می‌توان فقط از برخی لایه‌ها برای محاسبه فاصله استفاده شود تا سربار پردازشی کنترل شود. به‌طور کلی، تابع پیشنهادی با ادغام هم‌زمان سه مؤلفه اصلی شامل کیفیت یادگیری محلی، اهمیت ساختاری گره در گراف و میزان هم‌راستایی مدل‌ها، یک معیار امتیازدهی چندبعدی و تطبیقی ارائه می‌دهد. این ویژگی‌ها باعث می‌شود که گره‌های منتخب برای تجمع، علاوه بر برخورداری از عملکرد محلی مناسب، از نظر ساختاری و محتوایی نیز با سایر گره‌ها سازگار باشند. انتظار می‌رود استفاده از این تابع امتیازدهی در یادگیری فدرال غیرمتمرکز، به‌ویژه در محیط‌های پویا و ناهمگن، منجر به بهبود همگرایی، افزایش پایداری آموزش و کاهش اثر انحراف مشتریان شود. با الگوریتم حریم‌ناهیانه (۱)، مجموعه برتری به دست می‌آید.

این مجموعه برتری دقیقاً گره‌هایی هستند که در دور جاری در تجمع مشارکت می‌کنند. شکل (۱) مرحله اول روش پیشنهادی که ترکیب گراف‌های محلی دو گره در یک ملاقات است را نشان

جدید را از تجمع گراف محلی قبلی خود و گراف دریافت‌شده می‌سازند. وزن گره‌ها (w_i) و وزن یال‌ها (w_{ij}) به روش زیر محاسبه می‌شوند:

(۱) وزن گره i (w_i): ترکیبی از متغیرهای حجم داده (S_i) و دقت مدل محلی (A_i) که نرمال‌سازی شده و برای جلوگیری از تأثیر گره‌های ضعیف و پرت است که به مطابق رابطه (۱) تعریف می‌شود:

$$w_i = \left(\frac{S_i}{\sum_{j=1}^n S_j} \right) \times \left(\frac{1}{1 + e^{-(A_i - \theta)}} \right) \quad (1)$$

استفاده از تابع سیگموئید باعث می‌شود گره‌هایی که دقتی کمتر از حد آستانه (θ) دارند، وزن کمتری دریافت کنند.

(۲) وزن یال (w_{ij}): نشان‌دهنده شباهت مدل یادگیری عمیق دو گره i و j است. از دو معیار شباهت استفاده شده است:

الف) شباهت کسینوسی از بردارهای وزن مدل‌های مسطح‌شده استفاده می‌کند که ساده، سریع و پرکاربرد است [۲۰]. با استفاده از این روش مفهوم فاصله دو مدل که با استفاده از رابطه (۲) محاسبه می‌شود، به دست می‌آید.

$$d_{\text{cosine}}(M_i, M_j) = 1 - \frac{M_i \cdot M_j}{\|M_i\| \|M_j\|} \quad (2)$$

ب) همبستگی پیرسون یک معیار آماری کلاسیک است که قدرت و جهت رابطه خطی بین دو متغیر را اندازه‌گیری می‌کند. در یادگیری ماشینی، از این معیار به‌طور گسترده‌ای برای تحلیل شباهت‌های ساختاری بین مدل‌ها، مانند مقایسه بردارهای وزن یا فعال‌سازی‌های لایه‌های عصبی، استفاده می‌شود. ویژگی تغییرناپذیری نسبت به مقیاس خطی، آن را به ابزاری ارزشمند برای مقایسه مدل‌هایی تبدیل می‌کند [۲۱]. که در روابط (۳) و (۴) نحوه محاسبه همبستگی و فاصله بیان شده است.

$$r(M_i, M_j) = \frac{\sum_{i=1}^n (w_{1,i} - \bar{w}_1)(w_{2,i} - \bar{w}_2)}{\sqrt{\sum_{i=1}^n (w_{1,i} - \bar{w}_1)^2} \cdot \sqrt{\sum_{i=1}^n (w_{2,i} - \bar{w}_2)^2}} \quad (3)$$

¹ Dominance Score

می‌دهد. همان‌گونه که در شکل مشاهده می‌شود ابتدا گراف‌ها با هم ترکیب می‌شوند، سپس امتیاز گره‌ها محاسبه شده و گره‌های برتر با استفاده از محاسبه مجموعه برتری انتخاب می‌شوند. این گره‌ها در شکل (۱) با رنگ متفاوت نمایش داده شده‌اند.

تجمیع مدل شرکت داده نمی‌شوند. علت این امر آن است که گره‌های منتخب از نظر کیفیت و سطح مدل یادگیری محلی، حجم اطلاعات و اهمیت در گراف ارتباطی یکسان نیستند؛ بنابراین تخصیص سهم برابر به آن‌ها می‌تواند باعث کاهش کیفیت مدل نهایی شود. به منظور رفع این مشکل، در این پژوهش یک چارچوب تصمیم‌گیری چندمعیاره^۱ مبتنی بر ترکیب روش‌های فرآیند تحلیل سلسله مراتبی^۲ و WASPAS^۳ برای وزن‌دهی نهایی گره‌های مجموعه برتری به کار گرفته شده است. این روش باعث می‌شود هر گره بر اساس سطح شایستگی خود، سهمی متناسب در فرآیند تجمیع داشته باشد. در ابتدا باید معیارهایی که در فرآیند تجمیع تاثیرگذار هستند انتخاب شوند. برای هر گره عضو مجموعه برتری، سه معیار اصلی در نظر گرفته شده است:

الگوریتم (۱): الگوریتم مجموعه برتری حریصانه

```
Greedy_Dominating_Set(G(V,E)):
For each node v in V:
score[v] = compute_score(v)
nodes_list = sort_descending(V by score)
For node v in nodes_list:
if v is not yet covered:
D = D ∪ {v}
mark v and all its neighbors as 'covered'
Return D
```

۲-۲-۳ روش محاسبه وزن گره‌های منتخب

پس از تعیین مجموعه برتری گراف برای هر گره، تمام گره‌های موجود در این مجموعه به صورت مستقیم و با سهم برابر در فرآیند

(۱) دقت مدل محلی که نشان‌دهنده کیفیت عملکرد مدل آموزش‌دیده در آن گره است.

(۲) حجم داده محلی که بیانگر تعداد نمونه‌های آموزشی در اختیار گره است.

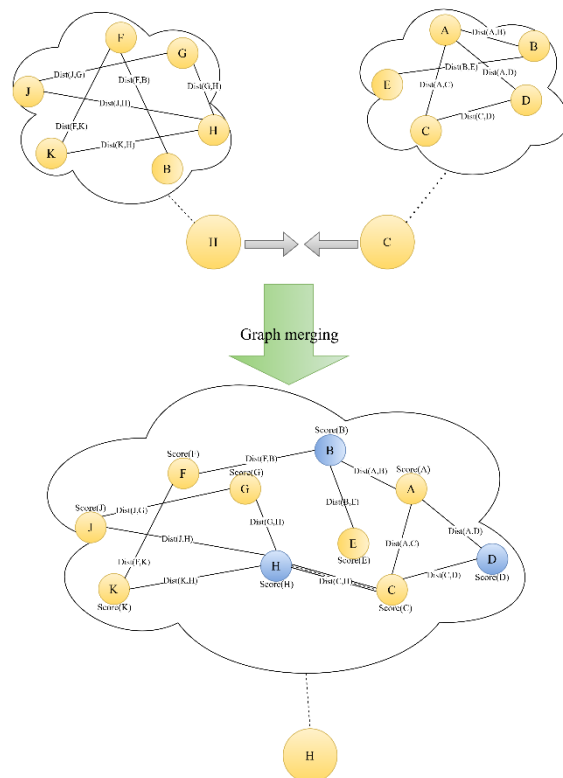
(۳) امتیاز گره در گراف محلی معیاری است که از ساختار گراف پویا و تعاملات گره با همسایگانش استخراج می‌شود و بیانگر اهمیت ساختاری آن گره است.

پس برای هر گره (i) در مجموعه برتری، یک بردار تصمیم مانند آنچه که در رابطه (۸) است، تعریف می‌شود.

$$x_i = [Accuracy_i, DataSize_i, Score_i] \quad (8)$$

مجموع این بردارها برای گره‌های مجموعه برتری، ماتریس تصمیم را تشکیل می‌دهد. در ادامه برای تعیین اهمیت نسبی معیارها، از روش فرآیند تحلیل سلسله مراتبی استفاده شده است. در این روش، ابتدا یک ماتریس مقایسه زوجی تشکیل می‌شود که در آن اهمیت هر معیار به صورت اولیه تعیین می‌شود. دقت مدل به عنوان معیار اول در نظر گرفته شده است. پس از آن حجم داده گره مشخص شده است و در نهایت امتیاز گره در گراف قرار گرفته است. ماتریس بیان شده در بالا به شکل رابطه (۹) تعریف می‌شود.

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \quad (9)$$



شکل (۱): ترکیب گراف‌های دو گره در یک ملاقات

³ Weighted Aggregated Sum Product Assessment

¹ Multi-Criteria Decision Making

² Analytic Hierarchy Process

پس از مشخص شدن سهم هر گره در فرآیند تجمیع حال در مرحله نهایی، وزن هر لایه از مدل جدید با استفاده از یک میانگین وزنی از وزن‌های مدل‌های اعضای مجموعه برتری به وسیله رابطه (۱۶) به دست می‌آید:

$$W_l^{(i)t+1} = \delta \times \sum_{i \in DS} \bar{Q}_i \cdot W_l^{(i)t} + (1 - \delta) \times W_l^{(i)t} \quad (16)$$

که در آن DS مجموعه برتری گراف محلی است و \bar{Q}_i وزن نهایی گره i است. همچنین $W_l^{(j)t}$ و $W_l^{(i)t}$ وزن فعلی لایه l در مدل گره (i) و گره (j) که گره‌ای است که عملیات تجمیع در آن انجام می‌شود. علاوه بر این، $\delta \in [0,1]$ یک پارامتر تنظیمی است تا بتوان شدت تغییرات در مدل‌های یک گره را مدیریت کرد. این روش باعث می‌شود گره‌هایی که دارای دقت بیشتر و اهمیت ساختاری بالاتری هستند، تأثیر بیشتری بر مدل نهایی داشته باشند، درحالی‌که تأثیر گره‌های ضعیف به‌طور طبیعی کاهش می‌یابد.

۴- ارزیابی

در این بخش به پیاده‌سازی روش پیشنهادی، شبیه‌سازی شرایط ارزیابی و نتایج حاصل از آن پرداخته می‌شود.

۴-۱ تنظیمات پیاده‌سازی، تحرک و توزیع داده‌ها

شبیه‌سازی در یک بستر سخت‌افزاری شامل کارت گرافیکی Nvidia RTX 3090 و پردازنده ایتل ۱۶ هسته‌ای انجام شده است. برای پیاده‌سازی روش پیشنهادی یک شبکه همتابه‌متای متحرک با مدل تحرک نقطه تصادفی مقصد در نظر گرفته شده است. در این مدل، هر گره به‌طور تصادفی یک نقطه مقصد در فضای شبیه‌سازی را انتخاب می‌کند، با سرعتی تصادفی (از بازه تعریف‌شده) به سمت آن حرکت می‌کند، پس از رسیدن به مقصد برای مدت زمان توقف^۳ مکث کرده و دوباره این فرآیند را تکرار می‌کند. مدل نقطه تصادفی مقصد به دلیل سادگی و واقع‌گرایی نسبی، یکی از پرکاربردترین مدل‌های تحرک در شبکه‌های متحرک و موردی متحرک به‌شمار می‌رود. به‌منظور بررسی تأثیر تعداد گره‌ها و ناهمگونی داده‌ها، تعداد گره‌های شبکه به‌صورت متغیر در

پس از نرمال‌سازی ستون‌های این ماتریس و محاسبه میانگین سطرها، بردار وزن معیارها مطابق رابطه (۱۰) به دست می‌آید.

$$w = [w_1, w_2, w_3] \quad (10)$$

که به ترتیب بیانگر اهمیت سه معیار دقت، حجم داده و امتیاز گره در گراف است. همچنین در این مرحله، با محاسبه شاخص سازگاری، سازگاری منطقی مقایسه‌های انجام‌شده بررسی می‌شود. در گام بعدی، برای مقابله با محدودیت‌های روش جمع وزنی ساده و افزایش دقت تصمیم‌گیری، از روش WASPAS که ترکیبی از دو مدل جمع وزنی و ضرب وزنی است، استفاده شده است. ابتدا ماتریس تصمیم نرمال‌سازی می‌شود. از آنجاکه تمامی معیارها از نوع سودمند (و نه هزینه) هستند، نرمال‌سازی با استفاده از رابطه (۱۱) انجام می‌شود.

$$\bar{x}_{ij} = \frac{x_{ij}}{\max(x_j)} \quad (11)$$

سپس دو امتیاز مستقل برای هر گره محاسبه می‌گردد که در ادامه با ترکیب کردن آن‌ها امتیاز و سهم هر گره در تجمیع مشخص می‌گردد.

محاسبه امتیاز جمع وزنی^۱ و امتیاز ضرب وزنی^۲ هر گره به‌وسیله روابط (۱۲) و (۱۳) انجام می‌شود.

$$Q_i^{(1)} = \sum_{j=1}^n w_j \cdot \bar{x}_{ij} \quad (12)$$

$$Q_i^{(2)} = \prod_{j=1}^n (\bar{x}_{ij})^{w_j} \quad (13)$$

در نهایت، امتیاز نهایی هر گره به‌صورت ترکیب خطی این دو مقدار با استفاده از رابطه (۱۴) به دست می‌آید.

$$Q_i = 0.5 \times Q_i^{(1)} + 0.5 \times Q_i^{(2)} \quad (14)$$

برای آنکه مجموع وزن‌ها برابر با یک شود و بتوان از آن‌ها به‌صورت ضرایب تجمیع استفاده کرد، نرمال‌سازی نهایی انجام می‌شود که در رابطه (۱۵) مشاهده می‌شود.

$$\bar{Q}_i = \frac{Q_i}{\sum_{k=1}^m Q_k} \quad (15)$$

در این رابطه، m تعداد گره‌های موجود در مجموعه برتری است.

³Stop time(pause time)

¹ Weighted Sum Model

² Weighted Product Model

سخت‌افزاری با کارت گرافیک NVIDIA RTX 3090 از ابرفردوسی استفاده شده است.

جدول (۲): دادگان‌های استفاده‌شده

CIFAR10	Fashion-MNIST	MNIST	جنبه
تصاویر اشیاء و حیوانات	تصاویر اقلام پوشاک و مد	تصاویر ارقام دست‌نویس	نوع داده
۱۰ کلاس	۱۰ کلاس	۱۰ کلاس	تعداد کلاس‌ها
۵۰,۰۰۰	۶۰,۰۰۰	۶۰,۰۰۰	تعداد نمونه‌های آموزش
۱۰۰۰۰	۱۰۰۰۰	۱۰۰۰۰	تعداد نمونه‌های آزمون
۳۲×۳۲ پیکسل	۲۸×۲۸ پیکسل	۲۸×۲۸ پیکسل	اندازه تصویر
رنگی	سیاه و سفید	سیاه و سفید	نوع تصویر
نسبتاً پیچیده	متوسط	ساده	سطح پیچیدگی

جدول (۳) نتایج مقایسه عملکرد روش پیشنهادی را با دو رویکرد دیگر در حوزه یادگیری فدرال، یعنی FedAvg و WAFL را نشان می‌دهد. به‌منظور بررسی تأثیر مقیاس‌پذیری بر کارایی روش‌ها، آزمایش‌ها در دو سناریوی مختلف با تعداد ۵۰ و ۱۰۰ مشتری و در طول ۱۰۰ دور، اجرا شده‌اند. همچنین توزیع داده‌های گره‌ها به‌صورت ناهمگون با یک حالت نسبتاً زیاد اما نه شدید ($\alpha = 0.5$) از ناهمگونی در نظر گرفته شده است. این تنظیمات به‌منظور بررسی رفتار روش پیشنهادی در شرایط مختلف تحرک گره‌ها و مقیاس‌پذیری سیستم انتخاب شده‌اند، زیرا افزایش تعداد مشتری‌ها، تحرک و جابجایی آن‌ها معمولاً منجر به چالش‌های بیشتری در همگرایی مدل سراسری می‌شود. معیار ارزیابی، دقت طبقه‌بندی بر روی دادگان آزمون است که در جدول ارائه گردیده است. علاوه بر نتایج عددی، در شکل (۲) نمودارهای مقایسه‌ای ترسیم شده‌اند. نتایج به‌دست‌آمده به‌وضوح نشان‌دهنده برتری روش پیشنهادی در تمامی پیکربندی‌های آزمایشی نسبت به دو الگوریتم FedAvg و WAFL است. این برتری در دادگان‌های با پیچیدگی بالاتر و غیرمستقل و غیرهم‌توزیع، مانند CIFAR10، مشهودتر است.

نظر گرفته شده است. داده‌های محلی هر گره به‌صورت غیرمستقل و غیرهم‌توزیع تولید شده‌اند. برای ایجاد این ناهمگونی از توزیع دیریکله استفاده شده است. توزیع دیریکله روشی برای شبیه‌سازی و مدل‌سازی ناهمگونی آماری داده‌ها در مطالعات یادگیری فدرال است. در این روش، متغیر α میزان ناهمگونی را کنترل می‌کند. مقادیر کوچک α منجر به توزیع‌های پراکنده و ناهمگون (هر مشتری دارای تعداد محدودی از کلاس‌ها) و مقادیر بزرگ α منجر به توزیع‌های یکنواخت و همگون (از همه کلاس‌ها) می‌شود. علاوه بر شبیه‌سازی، از این توزیع برای تخمین درجه ناهمگونی داده‌های محلی و تنظیم پارامترهای بهینه‌سازی به‌منظور مقابله با انحراف مدل نیز استفاده می‌شود [۲۲]. این تنظیم به ما امکان می‌دهد قدرت روش پیشنهادی را در شرایط واقعی و چالش‌برانگیز یادگیری فدرال ارزیابی کنیم. جدول (۱) تمامی فرآیندهای استفاده‌شده در تعریف مدل پیشنهادی را نشان می‌دهد.

جدول (۱): فرآیندها

مقدار	نماد	فرآیند
۵۰,۱۰۰	N	تعداد کل دستگاه‌ها
۱	E	تعداد دوره‌های آموزش محلی
۰/۰۰۱	γ	نرخ یادگیری
۰/۹	M	مومنتوم
۰/۴	λ	ضریب تعادل بین روش‌های محاسبه شباهت مدل‌ها
۰/۲,۰۳	θ	حد آستانه دقت
۰/۷	δ	ضریب تنظیم تغییرات مدل به نسبت دور قبلی
۰/۱,۰/۵,۱,۱۰۰	α	نرخ توزیع دیریکله

۴-۲ نتایج

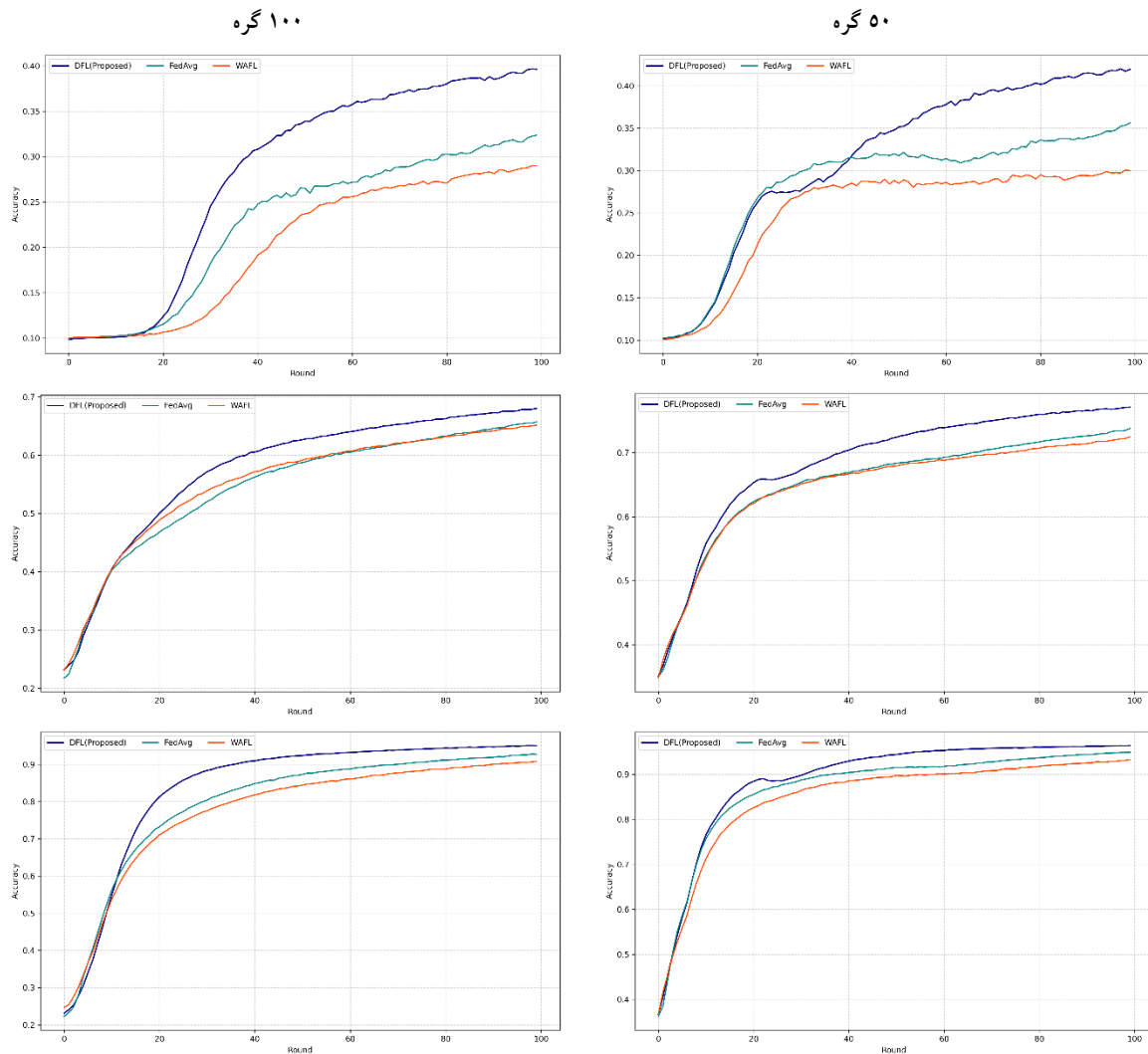
برای ارزیابی روش پیشنهادی از سه دادگان استاندارد و شاخص در حوزه بینایی ماشین شامل CIFAR10، Fashion-Mnist و MNIST استفاده شده است که جزئیات این دادگان‌ها در جدول (۲) آمده است [۲۳-۲۵]. همچنین برای پیاده‌سازی روش پیشنهادی و اجرای شبیه‌سازی در این مقیاس و تعداد مشتری از یک بستر

جدول (۳): نتایج به دست آمده توسط رویکرد پیشنهادی و رویکردهای FedAvg و WAFL

دادگان	تعداد مشتری	تعداد دور	دقت روش FedAvg	دقت روش WAFL	دقت روش پیشنهادی	بهبود (نسبت به FedAvg)	بهبود (نسبت به WAFL)
CIFAR10	۵۰	۱۰۰	۰/۳۶۱	۰/۳۰۳	۰/۴۲۴	۰/۰۶	۰/۱۲۱
CIFAR10	۱۰۰	۱۰۰	۰/۳۲۷	۰/۲۹۳	۰/۳۹۹	۰/۰۷۲	۰/۱۰۶
Fashion-MNIST	۵۰	۱۰۰	۰/۷۴۱	۰/۷۲۹	۰/۷۷۱	۰/۰۳	۰/۰۴۲
Fashion-MNIST	۱۰۰	۱۰۰	۰/۶۵۹	۰/۶۵۴	۰/۶۸۲	۰/۰۲۳	۰/۰۲۸
MNIST	۵۰	۱۰۰	۰/۹۵۰	۰/۹۳۶	۰/۹۶۴	۰/۰۱۴	۰/۰۲۸
MNIST	۱۰۰	۱۰۰	۰/۹۳۰	۰/۹۱۱	۰/۹۵۲	۰/۰۲۲	۰/۰۴۱

بهبودهای ۰/۰۲۳ و ۰/۰۲۸ در این سناریو نیز بیانگر کارایی روش پیشنهادی است. دادگان MNIST به عنوان یک مسئله پایه و نسبتاً ساده تر است، نتایج بر روی این دادگان نیز نشان دهنده برتری روش پیشنهادی است. در سناریوی ۵۰ مشتری، روش پیشنهادی به دقت ۰/۹۶۴ دست یافته که در مقایسه با دقت ۰/۹۵۰ روش FedAvg و دقت ۰/۹۳۶ روش WAFL، بهبودهای ۰/۰۱۴ و ۰/۰۲۸ را به همراه داشته است. در سناریوی دیگر با ۱۰۰ مشتری، دقت روش پیشنهادی ۰/۹۵۲ بوده که در مقابل دقت ۰/۹۳۰ روش FedAvg و دقت ۰/۹۱۱ روش WAFL، بهبودهای قابل توجه ۰/۰۲۲ و ۰/۰۴۱ را نشان می‌دهد. قابل ذکر است که میزان بهبود در این دادگان نسبت به دو مجموعه دیگر کمتر است که می‌تواند به دلیل اشباع شدن دقت در این مسئله ساده تر باشد. نتایج تجربی ارائه شده در جدول (۳) تأییدکننده این فرضیه است که روش پیشنهادی با بهره‌گیری از تئوری گراف و مکانیسم‌های پیشرفته انتخاب مشتریان مشارکت‌کننده در تجمیع، قادر به استخراج مؤثرتر دانش از داده‌های توزیع شده و کاهش اثرات منفی ناهمگونی بین مشتریان است. محاسبه میانگین بهبود در تمامی سناریوهای آزمایشی نشان می‌دهد که روش پیشنهادی به طور متوسط حدود ۳/۷ درصد (۰/۰۳۷) نسبت به FedAvg و ۶/۱ درصد (۰/۰۶۱) نسبت به WAFL بهبود دقت داشته است. این بهبود عملکرد را می‌توان به چند عامل نسبت داد.

دادگان CIFAR10 به دلیل تنوع کلاس‌ها و پیچیدگی ذاتی تصاویر، چالش برانگیزترین سناریو را ارائه می‌دهد. در سناریوی با ۵۰ مشتری، روش پیشنهادی با دستیابی به دقت ۰/۴۲۴، عملکرد بهتری نسبت به FedAvg با دقت ۰/۳۶۱ و WAFL با دقت ۰/۳۰۳ داشته است. میزان بهبود کارایی روش پیشنهادی نسبت به FedAvg معادل ۰/۰۶ (۶ درصد) و نسبت به WAFL برابر ۰/۱۲۱ (۱۲/۱ درصد) است. با افزایش تعداد مشتریان به ۱۰۰ مشتری، دقت هر سه روش با کاهش مواجه می‌شود که ناشی از چالش‌های ناهمگونی بیشتر داده‌ها در بین مشتریان است. با این حال، روش پیشنهادی با دقت ۰/۳۹۹، همچنان عملکرد برتر خود را حفظ کرده و بهبودهای ۰/۰۷۲ و ۰/۱۰۶ را به ترتیب نسبت به FedAvg و WAFL ثبت نموده است. این موضوع نشان دهنده مقاومت بیشتر روش پیشنهادی در برابر ناهمگونی بیشتر داده‌هاست. دادگان Fashion-MNIST که از نظر پیچیدگی رتبه پایین تری نسبت به دادگان CIFAR10 دارد، برای ارزیابی روش پیشنهادی انتخاب شده است. در این آزمون هم، روش پیشنهادی توانسته است دقت بالاتری را کسب کند. برای سناریوی ۵۰ مشتری، دقت روش پیشنهادی ۰/۷۷۱ است که در مقایسه با دقت ۰/۷۴۱ روش FedAvg و دقت ۰/۷۲۹ روش WAFL، بهبودهایی به ترتیب ۰/۰۳ و ۰/۰۴۲ را نشان می‌دهد. در سناریوی ۱۰۰ مشتری، اگرچه دقت کلی کاهش یافته، اما روش پیشنهادی با دقت ۰/۶۸۲ نسبت به رقبا ۰/۶۵۹ برای FedAvg و ۰/۶۵۴ برای WAFL عملکرد بهتری دارد.

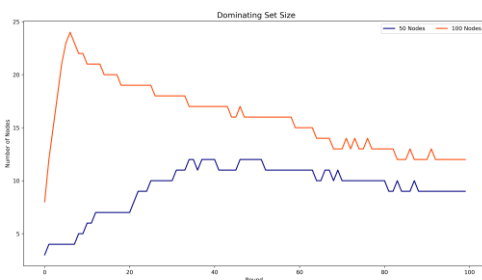


شکل (۲): نمودار دقت برای دادگان‌ها و تعداد گره‌های مختلف

است، که کارایی برتر روش پیشنهادی نشان‌دهنده مقیاس‌پذیری بهتر آن است. در مجموع، یافته‌های این بخش کارایی و اثربخشی رویکرد پیشنهادی را در دستیابی به دقت بالاتر، همگرایی سریع‌تر و پایداری بیشتر تا انتهای فرآیند آموزش برای مسائل یادگیری فدرال در سناریوهای متنوع و چالش‌برانگیز تأیید می‌کند. همچنین نمودار میله‌ای دقت روش‌ها و دادگان‌های مختلف در شکل (۳) آورده شده است.

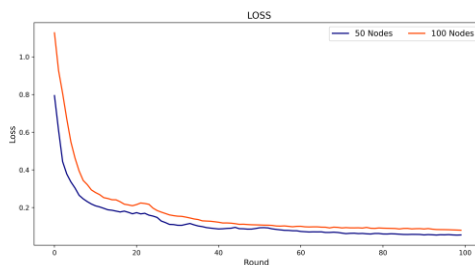
اولین و مهم‌ترین عامل، توانایی روش پیشنهادی در تعدیل به‌روزرسانی‌های محلی و جلوگیری از انحراف مدل ناشی از به‌روزرسانی‌های مشتریان ناکارآمد است. عامل دیگر، استفاده مؤثرتر از اطلاعات تجمعی و سوابق ملاقات‌ها در هر گره برای هدایت فرآیند یادگیری و عملیات تجمیع است. آخرین عامل هم مقاومت بالاتر در برابر نویز و داده‌های غیرکاربردی در مشتریان مختلف است. کاهش دقت همه روش‌ها با افزایش تعداد مشتریان (از ۵۰ به ۱۰۰) پدیده‌ای رایج و شناخته‌شده در یادگیری فدرال

اعضا افزایش می‌یابد و در ادامه با بهبود مدل‌ها و افزایش امتیاز گره‌ها، به صورت تدریجی تعداد گره‌های باکیفیت و امتیاز بالا به یک پایداری نسبی می‌رسند. این تعداد برای محیط با ۵۰ گره برابر با ۹ گره است که نسبتی در حدود ۱۸٪ تعداد کل گره‌ها را در برمی‌گیرد و برای محیط با ۱۰۰ گره برابر با ۱۲ گره است که این نسبت در حدود ۱۲٪ است. همگرایی مدل‌ها در رویکرد غیرمتمرکز در شبکه‌های متحرک به دلیل تحرک گره‌ها و تنوع داده‌ها یکی از چالش‌های دیگری است که باید به آن پرداخته شود.



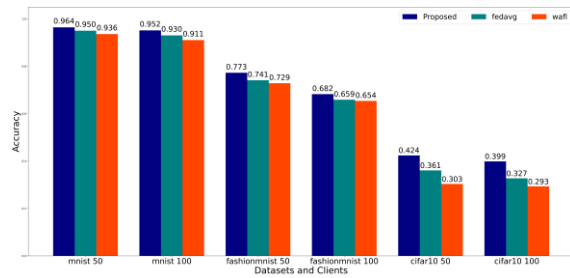
شکل (۵): نمودار تعداد اعضای مجموعه برتری گره‌ها

شکل (۶) نشان می‌دهد که تابع هزینه در دوره‌های اولیه با سرعت بالایی کاهش می‌یابد و سپس به صورت یکنواخت و آهسته به مقدار پایدار همگرا می‌شود. این روند تغییرات مقدار تابع هزینه بیانگر همگرایی سریع، کارایی و پایداری روش پیشنهادی است. علاوه بر این با بررسی میزان شباهت مدل‌های مشتری‌ها در دوره‌های مختلف هم می‌توان همگرایی مدل را ارزیابی کرد.



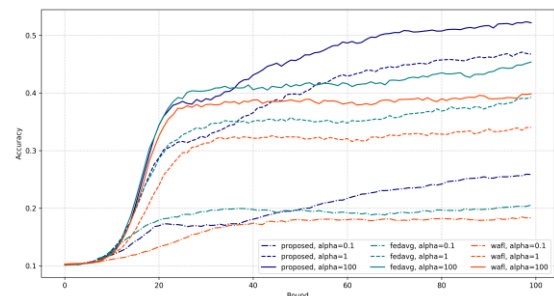
شکل (۶): نمودار همگرایی روش پیشنهادی

در شکل (۷) میزان شباهت مدل‌های مشتری‌ها در طول دورها مشاهده می‌شود. همان‌گونه که مشخص است برای محیط با ۵۰ مشتری، میزان شباهت کامل در دور ۵۳ آموزش اتفاق افتاده است که این میزان شباهت کامل در محیط با ۱۰۰ مشتری در دور ۲۴ آموزش مشاهده می‌شود.



شکل (۳): نمودار میله‌ای دقت دادگان‌ها

یکی از چالش‌هایی که در یک محیط متحرک یادگیری فدرال غیرمتمرکز ممکن است مشتریان با شدت بیشتری با آن مواجه شوند، میزان ناهمگونی داده‌های هر مشتری است. در شکل (۴) روند آموزش و تغییرات معیار دقت برای مقادیر مختلف ضریب توزیع دیریکله نشان داده شده است. همان‌گونه که مشاهده می‌شود هر چه میزان α بیشتر می‌شود، دقت و همگرایی روش پیشنهادی رشد بیشتری پیدا می‌کند. این رشد بیشتر به نسبت روش‌های دیگر به دلیل رویکرد انتخاب بهینه گره‌های مشارکت‌کننده در عملیات جمع است. علاوه بر این، تنظیم هدفمند وزن مشارکت مشتری‌ها هم نقش مهمی را در این بهبود دارد.



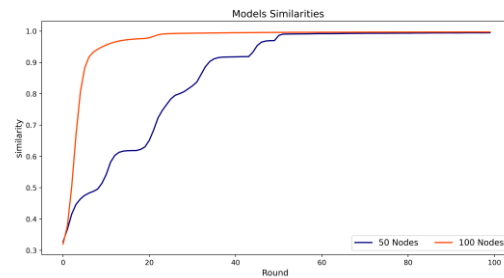
شکل (۴): نمودار تأثیر آبر روی دقت

همچنین با گذشت دوره‌های بیشتر و تکمیل ساختار گراف محلی، این انتخاب‌ها و وزندهی مشتری‌ها بالغ‌تر می‌شوند که می‌توان این در شکل (۵) نحوه تغییرات تعداد اعضای موجود در مجموعه برتری گره‌ها در دوره‌های مختلف برای مقادیر ۵۰ و ۱۰۰ گره نمایش داده شده است. در دوره‌های اولیه که تعداد ملاقات‌ها کمتر است نرخ تغییرات کمتر است، این مورد را به خصوص در نمودار مربوط به محیط ۵۰ گره به خوبی می‌توان مشاهده کرد، اما پس از گذار از دوره ابتدایی و افزایش دورها و ملاقات‌ها، تعداد گره‌های گراف محلی هر گره افزایش می‌یابد. با گذشت زمان در ابتدا تعداد

محدودیت‌هایی بر تعداد دوره‌های ارتباطی یا منابع محاسباتی وجود دارد، تبدیل می‌کند. این نتایج بیانگر توانایی بالای روش پیشنهادی در حل مشکلات ناشی از تحرک گره‌ها، مقیاس‌پذیری، انحراف مشتری و ناهمگونی داده در یادگیری فدرال کاملاً غیرمتمرکز است. مقایسه با FedAvg غیرمتمرکز و روش WAFL، به‌ویژه در مقیاس بزرگ (۵۰ و ۱۰۰ مشتری)، بهبود چشم‌گیری در دقت مدل ایجاد می‌کند، این در حالی است که این دو روش دچار افت در همگرایی می‌شوند. حفظ بهبود و روند افزایش آن با افزایش تعداد مشتری‌ها یکی از چالش‌هایی که در مورد روش پیشنهادی مطرح است پیچیدگی محاسباتی در محیط‌های خیلی پویا است که اگر مدل تحرک گره‌ها به گونه‌ای باشد که درجه گره‌ها در گراف محلی خیلی بالا باشد یعنی گره‌ها با نرخ بسیار بالایی گره‌های دیگر را مستقیم ملاقات کنند، سربار پردازشی و پیچیدگی محاسباتی افزایش پیدا می‌کند، بخش عمده این سربار پردازشی مربوط به محاسبه فاصله مدل‌های دو گره در گراف است. در نتیجه برای کارهای آینده تعمیم روش به سناریوهای واقعی‌تر با در نظر گرفتن مواردی چون خرابی تصادفی گره‌ها، هرس کردن گراف محلی، ملاحظات مربوط به فاصله دو گره از قبیل محدود کردن لایه‌ها در محاسبه فاصله و یا محدود کردن تعداد همسایگان برای محاسبه امتیاز گره در گراف محلی، تأخیرهای ارتباطی، ترکیب روش پیشنهادی با مکانیزم‌های رمزنگاری برای تضمین حفظ حریم خصوصی کامل در حین محاسبه شباهت مدل‌ها و ارزیابی چارچوب پیشنهادی روی وظایف پیچیده‌تر شامل داده‌های چندوجهی پیشنهاد می‌شود.

References

- [1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. Artif. Intell. Statist. (AISTATS), 2017, pp. 1273–1282.
- [2] P. Kairouz, H. B. McMahan et al., "Advances and Open Problems in federated Learning," Found. Trends Mach. Learn., vol. 14, no. 1, 2021.
- [3] J. Wang, Q. Liu, H. Liang, G. Joshi, and H. V. Poor, "Tackling the objective inconsistency problem in heterogeneous federated optimization," Adv. Neural Inf. Process. Syst., vol. 33, pp. 7611–7623, 2020.
- [4] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," Proc. Mach. Learn. Syst., vol. 2, pp. 429–450, 2020.
- [5] J. Qi, Q. Zhou, L. Lei, and K. Zheng, "Federated reinforcement learning: Techniques, applications, and open challenges," arXiv:2108.11887, 2021.
- [6] S. Kornblith, M. Norouzi, H. Lee, and G. Hinton, "Similarity of neural network representations revisited," in Proc. Int.



شکل (۷): نمودار شباهت مدل‌ها

۵- نتیجه‌گیری

این مقاله یک چارچوب کاملاً غیرمتمرکز یادگیری فدرال را در محیط‌های هم‌تابه‌ممتای پویا با داده‌های ناهمگون پیشنهاد می‌کند. روش پیشنهادی در هر دور ارتباطی، ابتدا با استفاده از مفهوم مجموعه برتری وزن‌دار و به‌صورت کاملاً محلی زیرمجموعه‌ای بهینه (۱۰٪–۲۰٪) از گره‌ها را بر مبنای معیارهای ترکیبی کیفیت مدل، حجم داده، سابقه مشارکت و شباهت عمیق مدل‌ها انتخاب می‌کند و سپس تجمیع وزن‌دار را تنها روی این گره‌های برتر انجام می‌دهد. شبیه‌سازی گسترده با مدل تحرک نقطه تصادفی مقصد و توزیع دیریکله روی دادگان‌های MNIST، Fashion-MNIST و CIFAR10 نشان می‌دهد که روش پیشنهادی در آزمایش‌های انجام‌شده بیانگر مقیاس‌پذیری برتر روش پیشنهادی در محیط‌های واقعی یادگیری فدرال با تعداد زیاد دستگاه‌های شرکت‌کننده است. این ویژگی‌ها، همراه با کاهش نوسانات در دوره‌های پایانی، روش پیشنهادی را به گزینه‌ای مناسب برای کاربردهای عملی که



- Conf. Mach. Learn. (ICML), 2019, pp. 3519–3529.
- [7] A. Karras et al., “Peer to peer federated learning: Towards decentralized machine learning on edge devices,” in Proc. SEEDA-CECNSM, 2022, pp. 1–9.
- [8] J. Qingzhe, “Decentralized P2P Federated Learning on Ad-hoc Like Networks with Non-IID Dataset,” M.S. thesis, Dept. Comput. Sci., Univ. Ottawa, Ottawa, ON, Canada, 2022.
- [9] Z. Qin et al., “BlockDFL: A blockchain-based fully decentralized peer-to-peer federated learning framework,” in Proc. ACM Web Conf. (WWW), 2024, pp. 2914–2925.
- [10] M. Mao et al., “SCAMA: A Smart-Contract-Driven Asynchronous Model Aggregation Framework for Decentralized Federated Learning,” in Proc. IEEE Int. Conf. High Perform. Smart Comput. (HPSC), 2025, pp. 102–107.
- [11] R. Chen et al., “FLock: Robust and Privacy-Preserving Federated Learning based on Practical Blockchain State Channels,” in Proc. ACM Web Conf. (WWW), 2025, pp. 884–895.
- [12] T. Shannon et al., “Exploring Heterogeneity in Federated Learning,” in Proc. Int. Conf. Wireless Artif. Intell. Comput. Syst. Appl., 2025, pp. 242–254.
- [13] S. Wang et al., “Towards Communication-Efficient Decentralized Federated Graph Learning over Non-IID Data,” arXiv:2509.08409, 2025.
- [14] H. Hegedűs, I. Hegedűs, A. Benczúr, and M. Jelasity, “Wireless Ad Hoc Federated Learning: A Fully Distributed Cooperative Machine Learning,” in Proc. IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW), 2019, pp. 148–153.
- [15] V. Hegiste, T. Legler, and M. Ruskowski, “Towards robust federated image classification: An empirical study of weight selection strategies in manufacturing,” in Proc. Int. Conf. Federated Learn. Technol. Appl. (FLTA), 2024, pp. 55–62.
- [16] Y. Dutta, A. Sikder, and B. Palit, “Federated Learning for Deforestation Detection: A Distributed Approach with Satellite Imagery,” arXiv:2509.13631, 2025.
- [17] B. Büyüktaş, G. Sumbul, and B. Demir, “A multi-modal federated learning framework for remote sensing image classification,” arXiv:2503.10262, 2025.
- [18] O. Odeyomi et al., “Online Decentralized Federated Multi-task Learning With Trustworthiness in Cyber-Physical Systems,” arXiv:2509.00992, 2025.
- [19] M. Pavan and M. Pelillo, “Dominant sets and pairwise clustering,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 1, pp. 167–172, Jan. 2007.
- [20] Y. Li et al., “ModelDiff: Testing Based DNN Similarity Comparison for Model Reuse Detection,” in Proc. ACM SIGSOFT Int. Symp. Softw. Test. Anal. (ISSTA), Jul. 2021, pp. 139–151.
- [21] M. Raghu et al., “SVCCA: Singular vector canonical correlation analysis for deep learning dynamics and interpretability,” Adv. Neural Inf. Process. Syst., vol. 30, 2017.
- [22] S. Arisdakessian, et al., “A Two-Level Dirichlet Framework for Heterogeneous Federated Network,” in IEEE Transactions on Network Science and Engineering, vol. 13, pp.1599-1615, 2026.
- [23] LeCun, Y., Cortes, C., & Burges, C. “The MNIST Database of Handwritten Digits”.1998.
- [24] Xiao, H., Rasul, K., & Vollgraf, R.. "Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms". arXiv:1708.07747, 2017.
- [25] Krizhevsky, A. “Learning Multiple Layers of Features from Tiny Images”. Technical Report, University of Toronto, CIFAR-10 and CIFAR-100 datasets, 2009.

A decentralized federated learning method based on dynamic graphs and multi-criteria decision making for mobile networks

Nima Karimi¹, Mohammad-Reza Pajooan^{2*}, Amir Jahangard Rafsanjani³

¹Phd candidate, Computer Engineering Department, Yazd University, Yazd, Iran

²Associate Professor, Computer Engineering Department, Yazd University, Yazd, Iran

³Assistant Professor, Computer Engineering Department, Yazd University, Yazd, Iran

Article Information

Original Research Paper

Received:

2025 December 29

Accepted:

2025 March 17

Keywords:

Decentralized federated learning, mobile peer-to-peer networks, graph dominance set, multi-criteria decision-making methods

Corresponding Author*:

pajooan@yazd.ac.ir

Abstract

Federated learning, as a new approach in distributed machine learning, enables training of artificial intelligence models without the need to transfer raw data. However, the classical federated learning architecture depends on a central server, which is not available in many real-world applications such as mobile networks and the Internet of Things. On the other hand, data heterogeneity, node mobility and communication limitations create fundamental challenges in implementing such systems. In this study, a new method for decentralized federated learning is proposed. In this approach each node updates the model solely based on its local communications and without dependence on a central server. A dynamic graph based on node encounters is constructed to model the communication structure and effective nodes in the aggregation process are selected using the computation of the graph dominance set. To accurately weight the selected models, a multi-criteria decision-making method is used that considers criteria such as accuracy, data volume, and node score. In addition, the similarity between models is used the graph combination process. The proposed method is evaluated on three benchmark datasets MNIST, FASHION-MNIST, and CIFAR10 and is compared with two baseline methods, peer-to-peer FedAvg and WAFL. The accuracy results of the models in two environments with 50 and 100 nodes yielding values of 0.964, 0.952, 0.771, 0.682, 0.424 and 0.399 respectively, indicate that the proposed method achieves superior performance. These findings demonstrate the capability of the proposed approach for application in dynamic and mobile environments with a large number of clients and without centralized infrastructure, where it can overcome the scalability issues and client drift problem in decentralized federated learning.

 : 10.22034/ABMIR.2026.24181.1213

E-ISSN: [2821-2037](https://doi.org/10.22034/ABMIR.2026.24181.1213)

/© 2026. Published by Yazd University This is an open access article under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).

