
بسم الله الرحمن الرحيم

دانشگاه صنعتی اصفهان
دانشکده مهندسی برق و کامپیوتر

بهبود کارایی الگوریتم یادگیری فدرال برای داده‌های غیرمستقل و غیریکنواخت با در نظر گرفتن میزان شباهت بین شبکه‌های عصبی در دستگاه‌های نهایی

پایان‌نامه کارشناسی ارشد مهندسی کامپیوتر - هوش مصنوعی و رباتیکز

علی بزرگ‌زاد

استادراهنما

دکتر امیر خورسندی

۱۴۰۳

فهرست مطالب

عنوان **صفحه**

فصل سوم: بررسی پیشینه روش‌های حل مشکل ناهمگنی آماری

چکیده

در این چکیده ...

کلمات کلیدی: یادگیری فدرال، یادگیری عمیق، ... .

فصل اول

مقدمه

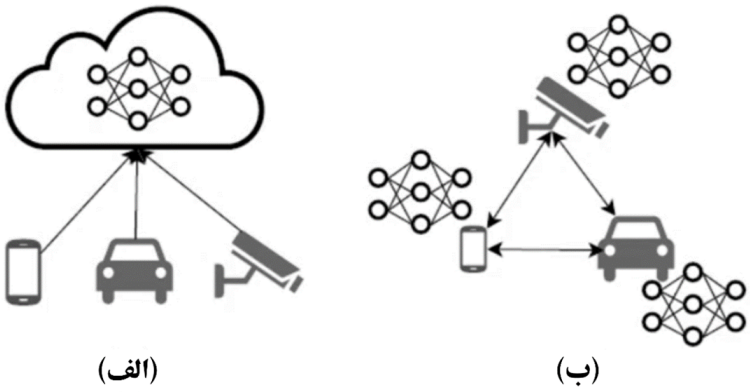
۱-۱ شناخت موضوع

رشد چشمگیر فناوری به همراه سهولت دسترسی به اینترنت در سال‌های اخیر باعث شده که بیشتر دستگاه‌های اطراف خود را متصل به اینترنت ببینیم. این دنیای جدید که به دنیای اینترنت اشیا[1] معروف است شامل خانه‌های هوشمند[2]، دستگاه‌های پوشیدنی[3]، خودروهای خودران و در صدر آن‌ها تلفن‌های هوشمند[4] است که همگی زندگی روزمره انسان را تغییر داده‌اند. استفاده از این سیستم‌ها همگی باعث تولید حجم قابل توجهی داده در طول روز می‌شوند که شرکت‌های بزرگ فناوری از این داده‌ها بهره برده و با استفاده از آن‌ها اقدام به انواع سرویس‌دهی به کاربران خود می‌نمایند.

با توجه به گسترش علم هوش مصنوعی و استفاده از روش‌های یادگیری ماشین، می‌توان از این حجم بسیار زیادِ داده تولید شده به نحو مطلوبی استفاده نمود و الگوریتم‌های مورد نظرمان جهت رسیدن به اهداف مختلف را بر روی آن‌ها اجرا کرد. حال برای مدیریت و اجرای الگوریتم‌های یادگیری، روش‌های مختلفی وجود دارد که به توضیح هر یک از آن‌ها خواهیم پرداخت.

---

[1] Internet of Things
[2] Smart Homes
[3] Wearable Devices
[4] Smart Phones



شکل ۱-۱: (الف) یادگیری متمرکز، (ب) یادگیری غیرمتمرکز [۲].

۱-۱-۱ یادگیری متمرکز

روش یادگیری متمرکز[1] که در اکثر سیستم‌های حال حاضر امروزی مورد استفاده قرار می‌گیرد به این نحو است که تمام گره‌ها[2] اطلاعات موجود خود را به صورت کامل به سمت سرویس‌دهنده ابری[3] ارسال می‌نمایند و سرویس‌دهنده ابری در حالی که تمام داده‌ها را در اختیار دارد اقدام به اجرای الگوریتم‌های مورد نظر می‌کند [۱]. در شکل ۱-۱ (الف) این روش به نمایش گذاشته شده است.

۱-۱-۲ یادگیری غیر متمرکز

در روش یادگیری غیر متمرکز[4] هر گره به صورت مجزا اقدام به اجرای الگوریتم‌های مورد نظر می‌کند و در واقع پس از اجرای چند مرحله از کد، اطلاعات به‌روز شده را با گره‌های همسایه به اشتراک می‌گذارد، این کار به قدری ادامه پیدا می‌کند تا همگی به مقدار تعیین شده همگرا شوند [۲]. در شکل ۱-۱ (ب) این روش به نمایش گذاشته شده است.

۱-۱-۳ یادگیری توزیع شده

روش یادگیری توزیع شده[5] به این نحو است که مدیریت کل سیستم و تمام داده‌ها در اختیار یک هسته مرکزی قرار دارد ولی به دلیل نیاز به توان پردازشی بالا، این هسته بار پردازشی را بین گره‌های موجود تقسیم می‌کند. در ابتدای راه یادگیری توزیع‌شده، فرض بر این بوده است که تمام گره‌ها توان پردازشی یکسانی داشته و داده‌ها به میزان مساوی بین گره‌ها پخش خواهند شد. در شکل ۱-۲ این روش به نمایش گذاشته شده است.

---
[1] Centralized Learning
[2] Nodes
[3] Cloud Server
[4] Decentralized Learning
[5] Distributed Learning

شکل ۱-۲: یادگیری توزیع شده [۲].

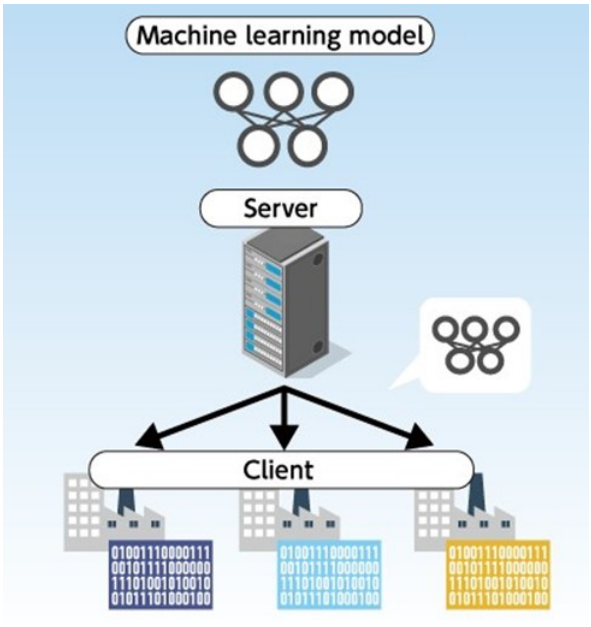
۱-۲ یادگیری فدرال

سیستم‌های متمرکز تا پیش از این، اکثر نیازهای مربوطه را برطرف می‌نمودند ولی در دنیای امروزی و با توجه به زیاد شدن هر روزه دستگاه‌های متصل، موارد دیگری نیز مورد توجه واقع شده است. هزینه‌های مرتبط با ارسال حجم زیاد داده از یک جهت، و افزایش اضطراب در مورد انتقال اطلاعات حساس و شخصی از جهت دیگر، محققان را به سمت بهره‌گیری از الگوریتم‌های غیرمتمرکز و توزیع شده در زمینه یادگیری ماشین هدایت کرده است. یکی از زیر مجموعه‌های روش‌های یادگیری توزیع‌شده، شاخه جدید و بسیار پراستفاده یادگیری فدرال بوده که بسیار مورد توجه قرار گرفته است.

در روش یادگیری فدرال، برخلاف رویکردهای متمرکز یادگیری ماشین، تجزیه و تحلیل داده‌ها به دستگاه‌های لبه[۱] یا سرویس‌گیرنده‌ها[۲] منتقل می‌شود. این روش، به عنوان یک جایگزین مطلوب و نوآورانه برای مدل‌سازی داده‌ها در محیط‌هایی با تعداد زیادی سرویس‌گیرنده معرفی شده است. در این چارچوب، به جای انتقال داده‌های اصلی، پارامترهای مدل‌های محلی در هر مرحله از فرآیند آموزش به سمت سرور منتقل می‌شوند، که این امر توانایی بهبود امنیت و کاهش هزینه‌های ارتباطی را فراهم می‌کند. در شکل ۱-۳ این روش به نمایش گذاشته شده است.

سرور در حقیقت نقش رهبری را ایفا می‌کند و با توجه به نوع داده‌ها، یک مدل شبکه عصبی[۳] ایجاد کرده و آن را به سمت کاربران ارسال می‌کند، حال کاربران با توجه به داده‌های خود شبکه را آموزش می‌دهند و بعد از چند بار تکرار، وزن‌های به‌روزرسانی شده را به سمت سرور بر می‌گردانند. همان‌طور که در شکل ۱-۳ مشاهده می‌شود، داده‌ها همگی در سمت کاربران قرار گرفته‌اند و به سمت سرور ارسال نمی‌شوند. عدم اجبار و محدودیت

[1] Edge Devices
[2] Clients
[3] Neural Network



شکل ۱-۳: یادگیری فدرال [۳].

در ارسال اطلاعات گره‌ها در یادگیری فدرال، به حفظ حریم شخصی کاربران کمک می‌کند [۴].

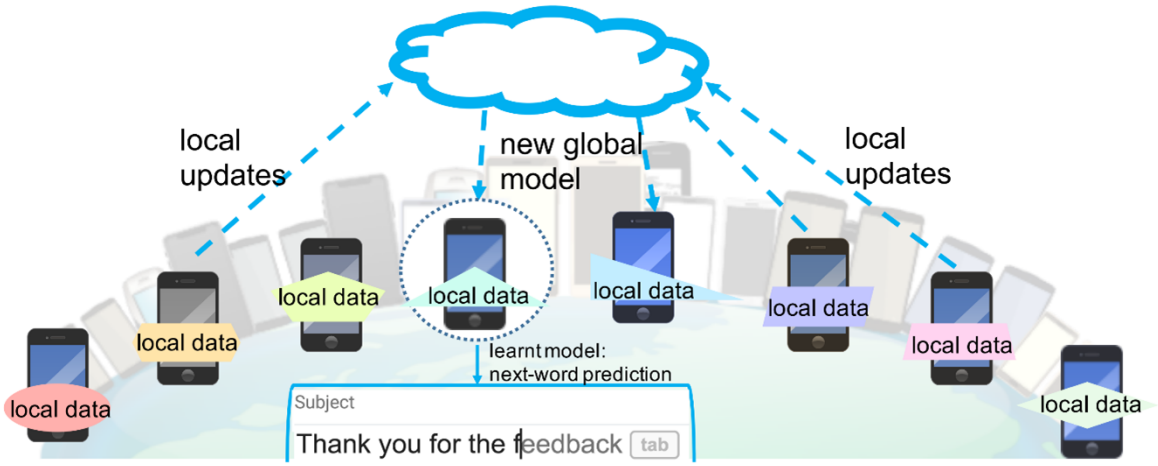
۱-۳ تاریخچه یادگیری فدرال

در ابتدای فصل بهار سال ۲۰۱۷ محققین گوگل (Google) طی یک مطلب کوتاه در وبلاگ هوش مصنوعی برای اولین بار موضوع یادگیری فدرال را تحت مطلبی با عنوان "یادگیری فدرال: یادگیری ماشین اشتراکی، بدون آموزش متمرکز داده‌ها" مطرح نمودند [۵]. در این مطلب به طور کوتاه Google Keyboard یا به اختصار Gboard معرفی شده و نحوه به کاری‌گیری یادگیری فدرال برای پیش‌بینی لغت بعدی را بیان می‌کند. یادگیری فدرال در این کاربرد نیاز به ارسال داده‌های کاربران به سمت سرور را حذف کرده است و به طور محلی مدل را به‌روزرسانی می‌کند. بنابراین، با بهره‌گیری از اطلاعات پنهان بسیار زیاد دستگاه‌ها در فرآیند مدل‌سازی، حریم شخصی سرویس‌گیرنده‌ها به نحوی بیشتر از پیش حفظ می‌شود. در شکل ۱-۴ نحوه استفاده از یادگیری فدرال در این برنامه به نمایش درآمده است.

۱-۴ کاربرد یادگیری فدرال

تکنولوژی نسل چهار صنعت[1]، دامنه ارتباطات نرم‌افزاری و سخت‌افزاری را در انواع مختلف سیستم‌ها گسترش داده است. این هماهنگی فناوری نرم‌افزار و سخت‌افزار، تبدیل به یک پدیده مهم در مجموعه‌ای از محیط‌های

[1] Industry 4.0



شکل ۱ - ۴: استفاده از یادگیری فدرال برای پیش‌بینی کلمه بعدی در Gboard [۶].

هوشمند و خودکار شده است. سنسورهای سابق که تنها مسئول اندازه‌گیری وضعیت‌ها بودند، جای خود را به دستگاه‌های هوشمند با قابلیت پردازش و برنامه‌ریزی داده‌ها سپرده‌اند. همچنین، گسترش ارتباطات در بستر اینترنت، امکان انتقال و تبادل داده‌ها بین انواع مختلف سیستم‌ها را ارائه کرده است. این پیشرفت‌ها منجر به کاهش نیاز به مرکزیت در تصمیم‌گیری و توسعه سیستم‌ها شده است و به وجود آورنده کنترل و نظارت پیشرفته و توزیع پردازش شده است. این ویژگی‌ها به همراه حجم بی‌سابقه داده، یادگیری فدرال را به یکی از بهترین روش‌های به‌کارگیری در توسعه سیستم‌های هوشمند تبدیل کرده است [۷]. در اینجا سه نمونه از کاربرد یادگیری فدرال را شرح خواهیم داد.

۱ - ۴ - ۱ یادگیری فدرال در شهر هوشمند

در یک شهر هوشمند[1]، اطلاعات جمع‌آوری شده از سنسورها، دستگاه‌ها و زیرساخت‌های مختلف، از جمله ترافیک، انرژی، پسماند و امنیت، به دلیل ارزش بالایی که دارند، به عنوان منبعی مهم برای بهبود عملکرد و کیفیت زندگی شهروندان محسوب می‌شوند. اما به همراه این ارزش‌ها، حفظ حریم خصوصی و امنیت اطلاعات شهروندان نیز امری بسیار حیاتی است. یادگیری فدرال به عنوان یک رویکرد نوین و مبتنی بر حفظ حریم خصوصی، در اینجا وارد عمل می‌شود.

این روش امکان پردازش داده‌های حساس مانند تصاویر، داده‌های محیطی و اطلاعات مکانی در محیط محلی و توزیع شده را فراهم می‌کند، به‌طوری‌که هر قسمت از شهر می‌تواند به صورت مستقل از سایر قسمت‌ها از این داده‌ها استفاده کند. این رویکرد امکان توسعه مدل‌های هوش مصنوعی و الگوریتم‌های بهبود عملکرد شهر هوشمند را با حفظ حریم خصوصی شهروندان فراهم می‌کند. به‌عنوان مثال، از طریق استفاده از یادگیری فدرال،

[1] Smart City

می‌توان بهبود در مدیریت ترافیک، بهینه‌سازی مصرف انرژی، کاهش آلودگی هوا و افزایش امنیت شهری را به دست آورد، در حالی‌که اطلاعات شخصی شهروندان محافظت می‌شود و از نگرانی‌های حریم خصوصی جلوگیری خواهد شد.

۱-۴-۲ یادگیری فدرال در بیمارستان

در یک بیمارستان، اطلاعات پزشکی بسیار حساس و مهم است که باید محفوظ و محرمانه نگهداری شود. اما در عین حال، استفاده از این داده‌ها برای بهبود خدمات بهداشتی و درمانی نیز بسیار ارزشمند است. در اینجا مفهوم یادگیری فدرال وارد عمل می‌شود. با استفاده از روش‌های یادگیری فدرال، بیمارستان می‌تواند از داده‌های پزشکی بیماران خود برای توسعه مدل‌هایی استفاده کند که پیش‌بینی میزان زمان بستری، بهبود در تشخیص بیماری‌ها و حتی افزایش بهره‌وری پزشکان را ایجاد می‌کنند، بدون اینکه این داده‌ها به‌طور مستقیم در اختیار یک مرکز جمع‌آوری اطلاعات واقع شوند.

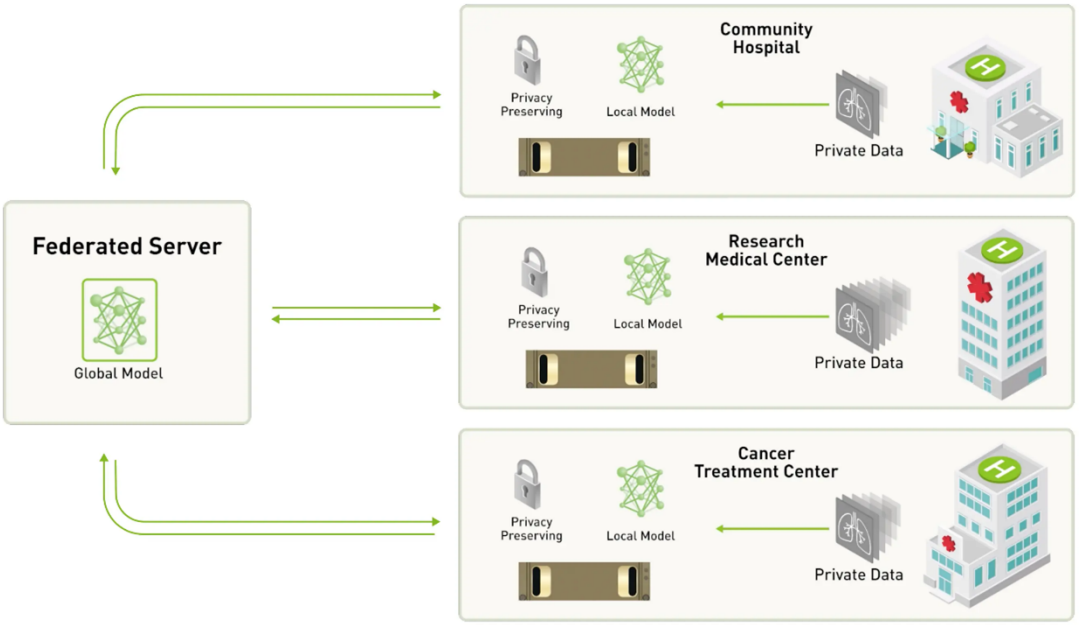
به عنوان مثال، با استفاده از یادگیری فدرال، مدل‌های هوش مصنوعی می‌توانند روی داده‌های محلی بیماران بیمارستان‌ها آموزش داده شوند تا بیماری‌های مختلف را شناسایی و تشخیص دهند، و اطلاعات مربوط به درمان‌های مؤثرتر را ارائه دهند، در حالی که اطلاعات حساس بیماران محافظت می‌شود. این روش به بیمارستان‌ها امکان می‌دهد که از داده‌های بیماران خود برای بهبود خدمات بهداشتی و درمانی استفاده کنند، در حالی که رعایت مقررات مربوط به حفظ حریم خصوصی و امنیت داده‌ها را به انجام رسانده‌اند. در شکل ۱-۵ یک نمونه استفاده از یادگیری فدرال در سازمان‌ها به نمایش در آمده است.

۱-۴-۳ یادگیری فدرال در فروشگاه برنامه‌های کاربردی موبایل

یک فروشگاه برنامه‌های کاربردی[1] موبایل را متصور شوید که به کاربران خود امکان می‌دهد برنامه‌های مختلف را دانلود و نصب کنند. این شرکت می‌خواهد با استفاده از داده‌های کاربران خود، الگوریتمی توسعه دهد که به طور دقیق‌تر بتواند پیشنهادات مربوط به برنامه‌هایی که کاربران ممکن است تمایل داشته باشند را ارائه کند.اگر این شرکت از روش‌های متمرکز استفاده کند، باید داده‌های حساس و شخصی کاربران را جمع‌آوری کند و برای آن‌ها تحلیل کند. این ممکن است باعث نگرانی‌های حریم خصوصی کاربران شود و از آن‌ها جلوگیری کند.

در حالی که با استفاده از یادگیری فدرال، این شرکت می‌تواند الگوریتم خود را بر روی داده‌های محلی هر تلفن هوشمند کاربر اجرا کند. به این ترتیب، هیچ داده‌ی حساسی به مرکز جمع‌آوری داده‌ها ارسال نمی‌شود و حریم خصوصی کاربران محفوظ می‌ماند. به عنوان مثال، اگر یک کاربر فقط به برنامه‌های موزیک علاقه‌مند

[1] App Store



شکل ۱ - ۵: یادگیری فدرال در یک بیمارستان [۸].

باشد، الگوریتم محلی در تلفن هوشمند او می‌تواند این الگو را تشخیص دهد و پیشنهادات مربوط به برنامه‌های موزیک را به او ارائه دهد، بدون این‌که داده‌های شخصی و حساس او به سرور شرکت ارسال شود. این روش به شرکت امکان می‌دهد از داده‌های کاربران خود برای بهبود خدمات خود استفاده کند، در حالی که حریم خصوصی آن‌ها را محافظت می‌کند.

۱ - ۵ دید کلی از روند موضوع و بیان هدف پژوهش

تکمیل این بخش پس از رسیدن به ساختار کلی پایان‌نامه (چون ممکنه در ادامه تغییر کنه)

چند جلمه کلیدی:

به دلیل پراکندگی همگرایی به کندی صورت می‌گیرد

روش جابجایی وزن‌ها بین کاربران نهایی در طول فرایند

چرا جابجایی تصادفی، جابجایی هوشمند بر اساس میزان شباهت

۱ - ۶ مروی بر روند ارائه مطالب پایان‌نامه

تست

فصل دوم

مفاهیم پایه در یادگیری فدرال

۲-۱ مقدمه

در جستجوی راه‌حل‌هایی برای یادگیری فدرال، لازم است به یک واقعیت مهم توجه کنیم که توزیع فرآیند آموزش بین افراد یا دستگاه‌های مختلف ممکن است به تداخل‌ها و مشکلاتی منجر شود. اگر این چالش‌ها را پیش از شروع فرآیند مدل‌سازی به‌خوبی در نظر نگیریم و راه‌حل‌های مشخصی برای آنها ارائه ندهیم، مدلی که در نهایت تولید می‌شود قطعاً با مشکلاتی از جمله دقت و کارایی مواجه خواهد شد. این مسئله، یکی از بزرگترین معظلاتی است که در مسیر یادگیری فدرال با آن روبرو می‌شویم و برای حل آن، نیازمند توجه دقیق و استفاده از روش‌های مختلف و نوآورانه هستیم.

در این فصل ابتدا چالش‌های موجود در یادگیری فدرال را رصد خواهیم کرد و سپس نگاه مقالات را در هر یک از آن‌ها به صورت کلی بررسی می‌کنیم. در ادامه، بیان ریاضی یادگیری فدرال را توضیح خواهیم داد و در نهایت به رویکردهای کلی و پایه‌ای در حل چالش‌ها اشاره خواهیم داشت.

۲-۲ چالش‌های موجود در یادگیری فدرال

با وجود مزیت‌های بسیار زیاد نسبت به روش‌های سنتی یادگیری ماشین، یادگیری فدرال به دلیل ساختار شبکه یادگیری با چالش‌های گوناگونی روبرو است. چالش‌های اصلی یادگیری فدرال عبارتند از:

۲-۲-۱ تبادل داده بین سرور و کاربران

تبادل داده بین سرور و کاربران به دلیل مشکلات پهنای باند و ارتباطات شبکه‌ای اصولا کار پر هزینه‌ای می‌باشد. یکی از دلایل اصلی پرهزینه بودن این ارتباطات، حجم بالای داده‌هایی است که باید بین دستگاه‌های کاربری و سرور منتقل شوند. معمولاً مشکلات ارتباطی به انتقال‌های بسیار زیاد به‌روزرسانی‌های گرادیان بین گره‌های محاسباتی نسبت داده می‌شوند. با افزایش تعداد پارامترها در مدل‌های پیشرفته، اندازه گرادیان‌ها نیز به طور متناسب بزرگ می‌شود [۹].

با این حال، تعداد زیادی از دستگاه‌های کاربر نهایی وجود دارند که در فرآیند آموزش مدل‌ها شرکت می‌کنند، که این موضوع می‌تواند هزینه‌های ارتباطات را به شدت افزایش دهد. علاوه بر این، در بسیاری از مواقع، همه دستگاه‌ها در هر چرخه از فرآیند آموزش شرکت نمی‌کنند، که این نیز به افزایش هزینه‌ها و پیچیدگی‌های مرتبط با انتقال داده‌ها منجر می‌شود.

۲-۲-۲ ناهمگنی‌های سیستمی

در دنیای یادگیری فدرال، دستگاه‌ها از نظر حافظه، توان محاسباتی و ارتباطات بسیار با یکدیگر متفاوت هستند. این تفاوت‌ها ممکن است از اختلافاتی مانند تفاوت در پردازنده، نوع حافظه، نوع اتصال شبکه و نیاز به انرژی ناشی شود. محدودیت‌های موجود در شبکه و سیستمی می‌توانند باعث ایجاد وضعیت‌هایی شوند که برخی از دستگاه‌ها در یک زمان معین در دسترس نباشند. برای مثال، اگر تعداد زیادی دستگاه همزمان درخواست ارسال داشته باشند، ممکن است برخی از آن‌ها به دلیل پهنای باند محدود یا محدودیت‌های سخت‌افزاری، قادر به ارسال درخواست نشوند. همچنین، ممکن است یک دستگاه فعال، به دلیل مشکلاتی مانند اختلالات در شبکه یا مصرف اضافی انرژی، از فرآیند یادگیری خارج شود.

این ویژگی‌های سیستمی، جزء اصلی چالش‌های یادگیری فدرال هستند و موجب افزایش تاخیر و اشکالات در سیستم می‌شوند. بنابراین، به منظور حل این مشکلات، روش‌های یادگیری فدرال باید قادر باشند تعداد دقیقی از دستگاه‌هایی که در فرآیند شرکت می‌کنند را پیش‌بینی کنند، همچنین باید در برابر دستگاه‌هایی که در حین عملیات با مشکل روبه‌رو شده‌اند مقاومت مناسبی داشته باشند [۶].

۲-۲-۳ ناهمگنی‌های آماری

طریقه تولید و جمع‌آوری داده‌ها بین دستگاه‌ها به شکل گوناگونی انجام می‌شود. این مجموعه داده‌ها اغلب مستقل از یکدیگر نیستند و ارتباطات و اتصالات میان آن‌ها وجود دارد. این الگوی تولید داده‌ها، با فرض استقلال و توزیع یکنواخت داده[1] (IID) در مسائل بهینه‌سازی متضاد است، که باعث ایجاد پیچیدگی در مدل‌سازی، تجزیه و تحلیل نظری و ارزیابی عملکرد راه‌حل‌ها می‌شود. در نتیجه، هرچند هدف نهایی یادگیری یک مدل سراسری است، اما روش‌های جایگزین مانند آموزش همزمان مدل‌های محلی جداگانه از طریق یادگیری چندوظیفه‌ای[2] و فرایادگیری[3]، به عنوان گزینه‌های جایگزین مطرح شده‌اند [۶].

۲-۲-۴ حریم شخصی

یکی از چالش‌های اساسی در یادگیری فدرال، حفظ حریم شخصی است که در این روش، داده‌های حساس و شخصی در اختیار بخش‌های مختلفی از شبکه قرار می‌گیرند. در این روش، دستگاه‌های محلی اطلاعاتی از کاربران جمع‌آوری و به سرور ارسال می‌کنند تا مدل‌های یادگیری مشترک را به‌روزرسانی کنند. این ارتباطات می‌توانند حاوی اطلاعات حساسی باشند که می‌توانند به راحتی به شناسایی فرد یا فرآیندهای حیاتی او منجر شوند.

یکی از مشکلات اساسی در اینجا این است که حتی با استفاده از روش‌های رمزنگاری و حفظ امنیت، اطلاعات معینی همچنان ممکن است به سرور ارسال شود که احتمالاً می‌تواند حریم شخصی را نقض کند. به‌طور خاص، اگر داده‌های حساس بدون رمزنگاری به سرور ارسال شوند یا اگر حتی اطلاعاتی که قابلیت شناسایی فرد را دارند به صورت رمزگذاری نشده ارسال شوند، حریم شخصی کاربران مورد تهدید قرار می‌گیرد.

۲-۳ نگاه مقالات مرتبط به چالش‌های موجود

۲-۳-۱ تبادل داده

با استفاده از فشرده‌سازی داده‌ها، می‌توان هزینه‌های ارتباطی را به طور قابل توجهی کاهش داد. دو روش جهت مدیریت هزینه‌های بالای ارتباطات در فرایند یادگیری فدرال مورد بررسی قرار گرفته است. این روش‌ها به فشرده‌سازی داده‌هایی که از دستگاه‌های کاربری به سرور مرکزی ارسال می‌شوند متمرکز شده‌اند. این فشرده‌سازی اطلاعات ارسالی به گونه‌ای است که حجم داده‌های ارسالی کم شده و در نتیجه، هزینه‌های مربوط به ارتباطات نیز کاهش یابد [۱۰].

---
[1] Independent and Identically Distributed
[2] Multi-Tasking
[3] Meta Learning

در روشی به نام PCFL[1] که یک رویکرد حفظ حریم خصوصی و البته بسیار کارآمد از نظر ارتباطی می‌باشد، شامل سه جزء کلیدی است که به ترتیب از فشرده‌سازی دوطرفه، فشرده‌سازی مکانی گرادیان‌ها و یک پروتکل حفظ حریم خصوصی که خود از تقسیم راز[2] و رمزنگاری همگام[3] برای محافظت از حریم خصوصی داده‌ها استفاده می‌کند، بهره گرفته است [۱۱].

۲-۳-۲ ناهمگنی سیستمی و آماری

برای مقابله با ناهمگنی سیستمی و آماری، روش‌هایی مانند تعادل در به‌روزرسانی مدل مطرح شده است. در این روش، وزن‌دهی به نمونه‌ها بر اساس میزان نیاز به آموزش در هر دستگاه صورت می‌گیرد. این کار باعث می‌شود که دستگاه‌های با حجم داده کمتر، وزن بیشتری در به‌روزرسانی مدل داشته باشند [۱۲]. در رویکرد دیگری به نام یادگیری فعال، دستگاه‌هایی که داده‌های خود را به سرور ارسال می‌کنند، فعالیت خود را به نحوی تنظیم می‌کنند که مدل از داده‌های مهم‌تر و کمتر دیده شده بیشتر یاد می‌گیرد. این روش می‌تواند به تعادل در آموزش مدل کمک کند و از ناهمگنی سیستمی جلوگیری کند [۱۰].

یک روش دیگر برای حل مشکل ناهمگنی سیستمی و آماری در یادگیری فدرال استفاده از رویکرد ترکیبی یا ترکیب روش‌های یادگیری محلی است. در این رویکرد، به جای استفاده از یک الگوریتم یادگیری مشترک برای تمام دستگاه‌ها، از چندین الگوریتم یادگیری محلی با تنوع مدل‌ها و تنظیمات مختلف استفاده می‌شود. سپس، اطلاعات مدل‌های محلی روی سرور یا گره مرکزی جمع‌آوری می‌شود و با استفاده از ترکیب این اطلاعات، یک مدل یادگیری مشترک بروزرسانی خواهد شد [۱۲].

۲-۳-۳ حریم شخصی

در روش حفظ حریم خصوصی تفاضلی[4] با افزودن نویز به نتایج محاسبات یا به داده‌های ورودی، اطمینان حاصل می‌کند که حضور یا عدم حضور یک نمونه داده خاص در مجموعه داده‌ها، تأثیر قابل توجهی بر خروجی محاسبات نداشته باشد. این روش به ویژه برای حفظ حریم خصوصی در یادگیری فدرال مفید است زیرا از افشای اطلاعات حساس از طریق پارامترهای مدل جلوگیری می‌کند [۱۳].

رویکرد رمزنگاری همگام امکان محاسبه روی داده‌های رمزنگاری شده را بدون نیاز به رمزگشایی آن‌ها فراهم می‌کند. این تکنیک به ویژه در یادگیری فدرال برای حفظ حریم خصوصی داده‌ها در حین انجام محاسبات مفید است زیرا نیاز به تغییر ماهیت داده نبوده و چون جابجایی در یادگیری فدرال بسیار زیاد رخ می‌دهد، این روش

---

[1] Privacy Communication efficient Federated Learning
[2] Secret Sharing
[3] Homomorphic Encryption
[4] Differential Privacy

بسیار کارا خواهد بود [۱۴].

۲-۴ بیان ریاضی یادگیری فدرال

برای ورود به مباحث ریاضی پایه در یادگیری فدرال، ابتدا باید به تعریف دقیق مسئله بهینه‌سازی مرکزی بپردازیم که در این حوزه مطرح می‌شود. در یادگیری فدرال، هدف اصلی یافتن مجموعه‌ای از پارامترهای مدل است که عملکرد کلی مدل را بر روی داده‌های توزیع‌شده بین تعداد زیادی دستگاه بهینه کند. هر دستگاه دارای داده‌های محلی است و یک تابع هزینه محلی بر اساس این داده‌ها برای آن دستگاه تعریف می‌شود. مسئله بهینه‌سازی کلی در یادگیری فدرال به دنبال کمینه کردن مجموع وزنی این توابع هزینه محلی است تا یک مدل جامع و یکپارچه حاصل شود.

ما یک طرح به‌روزرسانی همزمان را فرض می‌کنیم که به صورت دوره‌های ارتباطی پیش می‌رود. یک مجموعه ثابت از K مشتری وجود دارد که هر کدام یک مجموعه داده محلی ثابت دارند. در ابتدای هر دوره، یک کسر تصادفی C از مشتری‌ها انتخاب می‌شوند و سرور وضعیت فعلی پارامترهای مدل جهانی را به هر یک از این مشتری‌ها ارسال می‌کند. هر مشتری انتخاب شده سپس بر اساس وضعیت جهانی و مجموعه داده محلی خود محاسبات محلی را انجام می‌دهد و یک به‌روزرسانی به سرور ارسال می‌کند. سپس سرور این به‌روزرسانی‌ها را به وضعیت جهانی خود اعمال می‌کند و این فرآیند تکرار می‌شود [۱۵].

در حالی که ما بر اهداف شبکه عصبی غیرمحدب [1] تمرکز داریم، الگوریتمی که بررسی می‌کنیم قابل اعمال به هر هدف مجموع‌ـ‌متناهی [2] به صورت زیر است.

$$\min_{w\in\mathbb{R}^d} f(w) \quad \text{where} \quad f(w) \stackrel{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} f_i(w) \tag{۱-۲}$$

برای یک مسئله یادگیری ماشین، معمولاً $f_i(w) = \ell(x_i, y_i; w)$ در نظر گرفته می‌شود، به این معنی که این تابع نشان‌دهنده‌ی خطای پیش‌بینی بر روی نمونه (x_i, y_i) با استفاده از پارامترهای مدل w است. فرض می‌کنیم که داده‌ها بین K مشتری تقسیم شده‌اند، که در آن $\mathcal{P}_k$ مجموعه‌ای از نقاط داده مربوط به مشتری k است و $n_k = |\mathcal{P}_k|$ تعداد این نقاط داده را نشان می‌دهد. بنابراین، می‌توانیم فرمول ۲-۱ را به صورت زیر بازنویسی کنیم:

$$f(w) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(w) \quad \text{where} \quad F_k(w) = \frac{1}{n_k}\sum_{i\in\mathcal{P}_k} f_i(w) \tag{۲-۲}$$

اگر مجموعه $\mathcal{P}_k$ با توزیع یکنواخت تصادفی از مثال‌های آموزشی بین مشتری‌ها تشکیل شده باشد، در آن

[1] Non-Convex
[2] Finite-Sum

صورت $\mathbb{E}_{\mathcal{P}_k}[F_k(w)] = f(w)$ خواهد بود، که در اینجا امید ریاضی بر روی مجموعه مثال‌های اختصاص داده شده به یک مشتری ثابت گرفته می‌شود. این همان فرض IID (استقلال و توزیع یکسان) است که عموماً توسط الگوریتم‌های بهینه‌سازی توزیع‌شده استفاده می‌شود، در این‌جا ما حالتی را که این فرض برقرار نیست (یعنی F_k می‌تواند تقریباً به هر میزانی از f فاصله داشته باشد) به عنوان حالت Non-IID (غیرمستقل و غیریکنواخت) می‌شناسیم [۱۵].

۲-۵ رویکردهای کلی و پایه‌ای در حل چالش‌ها

روش‌های بهینه‌سازی توزیع‌شده معمولاً برای حل مسائل بهینه‌سازی در سیستم‌هایی با شبکه‌های محاسباتی بزرگ و توزیع‌شده استفاده می‌شوند. این روش‌ها بر مبنای تقسیم مسئله بهینه‌سازی به زیرمسائل کوچک‌تر و حل آن‌ها در گره‌های مختلف شبکه استوارند. در این روش‌ها، اغلب فرض می‌شود که داده‌ها به صورت همگن و یکپارچه در سراسر شبکه توزیع شده‌اند و گره‌ها می‌توانند به راحتی با یکدیگر ارتباط برقرار کنند.

فرضیات مطرح شده در یادگیری فدرال به ندرت برقرار است، زیرا در یادگیری فدرال داده‌ها به صورت محلی و ناهمگن در دستگاه‌های مختلف قرار دارند و ارتباطات بین دستگاه‌ها ممکن است محدود و نامنظم باشد. بنابراین روش‌ها و رویکردهای لازم جهت حل این چالش‌ها متفاوت از مسائل بهینه‌سازی توزیع شده هستند. حال سعی می‌کنیم دو رویکرد پایه‌ای برای مسائل یادگیری فدرال را مطرح نماییم.

۲-۵-۱ به‌روزرسانی محلی و میانگین‌گیری در سرور

یکی از روش‌های اصلی و پرکاربرد در یادگیری فدرال روش میانگین‌گیری فدرال[1] (FedAvg) است که توسط محققان گوگل در سال ۲۰۱۷ معرفی شد. این الگوریتم به منظور بهینه‌سازی مدل‌های یادگیری ماشین در یک محیط توزیع‌شده طراحی شده است، جایی که داده‌ها به صورت محلی در دستگاه‌های کاربران باقی می‌مانند و تنها به‌روزرسانی‌های مدل به اشتراک گذاشته می‌شوند. رویکرد اصلی FedAvg بر مبنای ترکیب به‌روزرسانی‌های محلی از دستگاه‌های مختلف به یک مدل جهانی استوار است.

یکی از مزایای اصلی FedAvg این است که به طور موثری با چالش ناهمگنی داده‌ها مقابله می‌کند. در یادگیری فدرال، داده‌های موجود در دستگاه‌های مختلف ممکن است توزیع‌های متفاوتی داشته باشند. این ناهمگنی می‌تواند به دلیل تفاوت در رفتار کاربران یا حتی محیط‌های مختلف جمع‌آوری داده باشد. میانگین‌گیری وزنی در FedAvg به مدل کمک می‌کند تا به‌روزرسانی‌های مختلف را به گونه‌ای ترکیب کند که این ناهمگنی‌ها را در نظر بگیرد. به عبارت دیگر، اگر یک دستگاه داده‌های بیشتری داشته باشد، تأثیر بیشتری بر مدل نهایی خواهد

[1] Federated Averaging

داشت. این رویکرد باعث می‌شود که مدل فدرال به تعادل بهتری در یادگیری از داده‌های ناهمگن برسد و کارایی بالاتری داشته باشد. این ویژگی به خصوص در کاربردهایی که کاربران متنوع و داده‌های متنوعی دارند، بسیار مفید است و می‌تواند به بهبود عملکرد مدل در شرایط واقعی کمک کند.

علاوه بر این، FedAvg به کاهش نیاز به ارتباطات مکرر بین دستگاه‌ها و سرور مرکزی کمک می‌کند. در بسیاری از روش‌های بهینه‌سازی توزیع‌شده، نیاز است که دستگاه‌ها به طور مکرر با سرور مرکزی ارتباط برقرار کنند تا به‌روزرسانی‌های خود را ارسال کنند. اما در FedAvg دستگاه‌ها می‌توانند چندین مرحله از بهینه‌سازی را به صورت محلی انجام دهند و سپس تنها به‌روزرسانی نهایی را ارسال کنند. این کاهش در نیاز به ارتباطات نه تنها باعث کاهش پهنای باند مورد نیاز می‌شود، بلکه به حفظ حریم خصوصی کاربران نیز کمک می‌کند، زیرا داده‌ها هرگز از دستگاه‌های محلی خارج نمی‌شوند. بررسی‌ها نشان داده‌اند که متناسب با اندازه داده‌ها پس از رسیدن به تعداد معینی از گره‌ها، اضافه کردن گره‌های بیشتر تأثیری در کاهش هزینه‌های ارتباطی نخواهد داشت. در چنین شرایطی، تمرکز بر افزایش توان محاسباتی محلی یا تعداد مراحل آموزش محلی می‌تواند موجب تسریع فرایند آموزش شود [۱۵].

موفقیت‌های اخیر در کاربردهای یادگیری عمیق تقریباً به‌طور انحصاری به استفاده از انواع الگوریتم نزول گرادیان تصادفی[1] (SGD) برای بهینه‌سازی متکی بوده‌اند. در واقع، بسیاری از پیشرفت‌ها به تنظیم مدل و بهینه‌سازی تابع خطا با روش‌های ساده گرادیان مربوط می‌شود. بنابراین، طبیعی است که ما الگوریتم‌های بهینه‌سازی فدرال را با شروع از SGD بسازیم.

الگوریتم SGD می‌تواند به سادگی در بهینه‌سازی فدرال استفاده شود، به این صورت که در هر دور ارتباط، گرادیان‌ها بر اساس داده‌های یک مشتری تصادفی انتخاب شده، محاسبه شوند. این رویکرد از نظر محاسباتی کارآمد است، اما نیازمند تعداد بسیار زیادی از دورهای آموزش برای تولید مدل‌های خوب است. برای مثال حتی با استفاده از رویکرد پیشرفته‌ای مانند نرمال‌سازی دسته‌ای[2]، برای آموزش دیتاست معروف MNIST (دیتاستی جهت دسته‌بندی اعداد دستنویس بین صفر تا نه) با دسته‌های کوچکی به اندازه ۶۰ به ۵۰۰۰۰ دور آموزش جهت رسیدن به مدل مطلوب نیاز می‌باشد [۱۶].

در تنظیمات فدرال، مشارکت تعداد زیادتری از مشتریان هزینه‌ای آنچنان بیشتری در زمان واقعی ندارد زیرا همه کاربران می‌توانند به صورت همزمان اقدام به آموزش مدل محلی کنند، بنابراین برای خط مبنای خود از SGD همزمان با دسته‌های بزرگ استفاده می‌کنیم. برای اعمال این رویکرد در تنظیمات فدرال، ما در هر دور یک کسر C از مشتریان را انتخاب می‌کنیم و گرادیان خطا روی تمام داده‌های نگهداری شده توسط این مشتریان را

[1] Stochastic Gradient Descent
[2] Batch Normalization

محاسبه می‌کنیم. بنابراین C اندازه دسته کلی را کنترل می‌کند، به‌طوری که $C = 1$ معادل با نزول گرادیان یک دسته کامل است. حال این الگوریتم خط مبنا را FederatedSGD یا FedSGD می‌نامیم.

یک پیاده‌سازی معمول از FedSGD با $C = 1$ و نرخ یادگیری ثابت η به این صورت است که هر گره k، گرادیان $g_k = \nabla F_k(w_t)$ که میانگین گرادیان روی داده‌های محلی در مدل فعلی w_t است را محاسبه می‌کند و سرور مرکزی این گرادیان‌ها را جمع‌آوری کرده و به‌روزرسانی $w_{t+1} \leftarrow w_t - \eta \sum_{k=1}^{K} \frac{n_k}{n} g_k$ را انجام می‌دهد، در حالی که $\sum_{k=1}^{K} \frac{n_k}{n} g_k = \nabla f(w_t)$ خواهد بود. یک به‌روزرسانی معادل به این صورت است که برای هر گره عبارت $\forall k, w_{t+1}^k \leftarrow w_t - \eta g_k$ محاسبه و سپس $w_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^k$ انجام شود.

در نتیجه، هر گره به صورت محلی یک گام گرادیان نزولی را روی مدل فعلی با استفاده از داده‌های محلی خود طی می‌کند و سپس سرور میانگین وزنی مدل‌های حاصل را محاسبه می‌کند. وقتی که الگوریتم به این صورت نوشته شود، می‌توانیم با تکرار به‌روزرسانی محلی $w^k \leftarrow w^k - \eta \nabla F_k(w^k)$، چندین بار قبل از مرحله میانگین‌گیری، محاسبات بیشتری به هر گره اضافه کنیم. در نهایت این رویکرد جدید را FederatedAveraging (FedAvg) می‌نامیم.

میزان محاسبات توسط سه پارامتر کلیدی کنترل می‌شود: C، کسر گره‌هایی که در هر مرحله محاسبات انجام می‌دهند؛ E، تعداد مراحل آموزشی که هر گره در هر دور روی مجموعه داده محلی خود انجام می‌دهد؛ و B، اندازه دسته محلی که برای به‌روزرسانی‌های هر گره استفاده می‌شود. در اینجا $B = \infty$ را می‌نویسیم تا نشان دهیم که کل مجموعه داده محلی به عنوان یک دسته واحد در نظر گرفته می‌شود. بنابراین، به عنوان یک نمونه از این الگوریتم گسترده شده جدید، می‌توانیم $B = \infty$ و $E = 1$ را انتخاب کنیم که در این حالت دقیقاً با FedSGD برابر خواهد شد. همچنین برای یک گره با n_k نمونه محلی، تعداد به‌روزرسانی‌های محلی در هر دور با $u_k = E\frac{n_k}{B}$ نمایش داده می‌شود [۱۵].

۲-۵-۲ بهینه‌سازی FedProx

روش FedProx به بررسی چالش‌های یادگیری فدرال در بسترهای ناهمگن می‌پردازد. این روش با ایجاد تغییرات جزئی در روش موجود FedAvg، به بهبود پایداری و دقت در شبکه‌های ناهمگن کمک می‌کند. این تغییرات شامل اضافه کردن یک عبارت نزدیک مبدا[1] به تابع هدف است که به صورت اصولی به سرور کمک می‌کند تا ناهمگنی را مدیریت کند.

[1] Proximal Term

فرمول هدف FedProx به صورت زیر تعریف می‌شود:

$$\min_{w} f(w) = \min_{w} \sum_{k=1}^{K} \frac{n_k}{n} \left(F_k(w) + \frac{\mu}{2} \|w^t - w_k^t\|^2 \right) \quad (۳-۲)$$

در فرمول ۲-۳ بخش $\frac{\mu}{2}\|w^t - w_k^t\|^2$، همان عبارت نزدیک مبدا است که به تابع هدف اضافه شده است. همچنین μ، یک پارامتر تنظیم برای این عبارت به حساب می‌آید و در نهایت w_k^t وزن‌های مدل محلی دستگاه k در تکرار t است.

حال با توجه به فرمول ۲-۳، به‌روزرسانی وزن‌ها به شکل زیر تغییر پیدا خواهد کرد و بخش $\mu(w^t - w_k^t)$، گرادیان عبارت نزدیک مبدا است.

$$w^{t+1} = w^t - \eta(\nabla F_k(w^t) + \mu(w^t - w_k^t))$$

بنابراین، به‌روزرسانی‌های محلی در هر گام با به‌روزرسانی سراسری مرحله قبل مرتبط هستند. عبارت نزدیک مبدا به عنوان یک مکانیزم منظم‌کننده[1] عمل می‌کند که تفاوت‌های بین وزن‌های جهانی w و وزن‌های محلی w_k^t را کاهش می‌دهد. این ترم به کاهش تاثیرات منفی ناهمگنی سیستم‌ها و داده‌ها کمک می‌کند و باعث پایداری بیشتر در فرآیند همگرایی می‌شود [۱۷].

[1] Regularization

فصل سوم

بررسی پیشینه روش‌های حل مشکل ناهمگنی آماری

۳-۱ مقدمه

همان‌طور که در فصل گذشته اشاره شد، یکی از مهم‌ترین مشکلات در حوزه یادگیری فدرال، مسئله داده‌های غیرمستقل و غیریکنواخت (non-IID) است که منجر به بروز چالش‌ها و ناهمگنی‌های آماری می‌شود. این مشکل باعث می‌شود که مدل‌های یادگیری نتوانند به خوبی از داده‌های توزیع‌شده استفاده کنند و کارایی مطلوبی داشته باشند. به دلیل اهمیت بالای این موضوع، محققان بسیاری تلاش‌های گسترده‌ای برای حل این مشکل انجام داده‌اند.

مبحث اصلی این پایان‌نامه نیز به طور دقیق به همین مسئله اشاره دارد و به دنبال یافتن راه‌حلی مؤثر برای مقابله با داده‌های non-IID است. در ادامه، به صورت خلاصه به بررسی راه‌حل‌هایی که تاکنون برای حل این مشکل مطرح شده‌اند، خواهیم پرداخت تا تصویر جامعی از تلاش‌های انجام شده در این زمینه ارائه دهیم. همچنین باید توجه داشت که هر یک از این راه‌حل‌ها نقاط قوت و ضعف خاص خود را دارند و بسته به شرایط و نوع داده‌ها، می‌توانند نتایج متفاوتی را به همراه داشته باشند. بررسی دقیق این راه‌حل‌ها و ارزیابی کارایی آن‌ها می‌تواند به بهبود سیستم‌های یادگیری فدرال و غلبه بر مشکلات مرتبط با داده‌های غیرمستقل و غیر یکنواخت کمک شایانی کند.

۳-۲ نگرش برپایه داده

۳-۲-۱ اشتراک‌گذاری داده

مشکل اصلی الگوریتم FedAvg در مواجهه با داده‌های غیرمستقل و غیریکنواخت، تفاوت وزن‌های اولیه در شروع فرآیند آموزش است. این تفاوت‌ها می‌توانند باعث شوند که مدل‌های محلی در هر گره به طور قابل توجهی متفاوت از یکدیگر باشند، که در نتیجه منجر به مشکلات همگرایی و کاهش کارایی مدل نهایی می‌شود.

برای رفع این مشکل، روشی پیشنهاد شده است که در آن ابتدا سرور مرکزی مقدار کمی از داده‌ها را به صورت محلی آموزش می‌دهد. در این مرحله، سرور مرکزی با استفاده از این داده‌ها، یک مدل اولیه را آموزش داده و وزن‌های اولیه آن را تنظیم می‌کند. سپس، این وزن‌های اولیه به همراه داده‌های آموزش دیده شده به تمامی کاربران ارسال می‌شود. این اقدام باعث می‌شود که تمام کاربران در ابتدای فرآیند آموزش با مجموعه‌ای از داده‌های مشترک و وزن‌های اولیه مشابه روبه‌رو شوند.

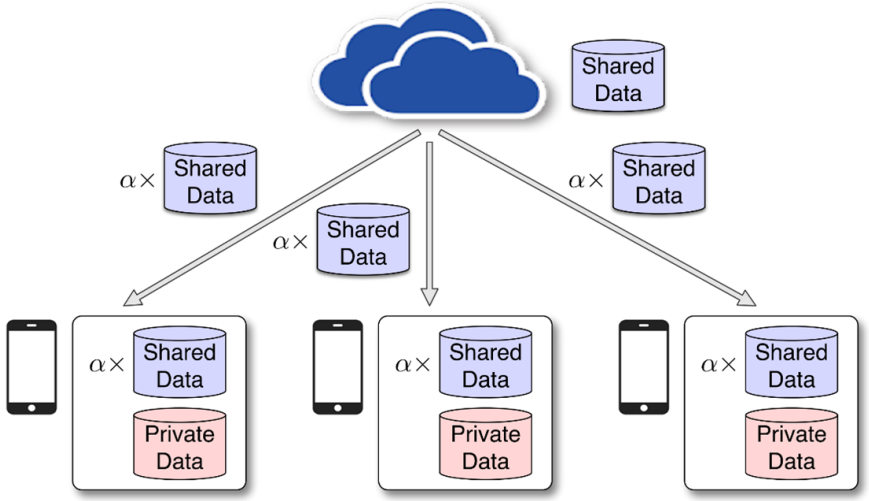
نقطه قوت این روش در این است که به دلیل انجام این عملیات تنها در آغاز فرآیند آموزش، هزینه زیادی به شبکه تحمیل نمی‌شود. در واقع، انتقال داده‌ها و وزن‌ها فقط در ابتدا انجام شده و پس از آن کاربران به صورت مستقل به آموزش مدل‌های محلی خود ادامه می‌دهند. این اقدام منجر به کاهش اختلافات ناشی از ناهمگنی داده‌ها شده و فرآیند همگرایی مدل نهایی سریع‌تر و با دقت بیشتری انجام می‌شود [۱۸].

در شکل ۳-۱، نحوه اجرای این روش و مراحل مختلف آن به تصویر کشیده شده است. این تصویر نشان می‌دهد که چگونه سرور مرکزی ابتدا داده‌های کمی را آموزش می‌دهد، وزن‌های اولیه را تنظیم می‌کند و سپس این وزن‌ها و داده‌ها را به کاربران ارسال می‌کند تا فرآیند آموزش محلی با یک نقطه شروع مشترک برای همه کاربران آغاز شود.

یکی دیگر از روش‌های مطرح شده در زمینه یادگیری فدرال به این صورت است که کاربران بتوانند نتایج آموزش تعدادی داده اشتراکی را با یکدیگر به اشتراک بگذارند و از نتایج دیگر کاربران بر روی این داده‌های اشتراکی مطلع شوند. در این روش، کاربران نتایج به‌دست آمده از آموزش داده‌های مشترک را با هم مبادله می‌کنند، که این کار منجر به بهبود عملکرد مدل‌های محلی و در نهایت مدل سراسری می‌شود [۱۹].

براساس بررسی‌های انجام شده، مثلاً در مجموعه داده CIFAR-10، اگر حدود ۵ درصد از داده‌ها به صورت اشتراکی در اختیار کاربران قرار گیرد، دقت مدل تا حدود ۳۰ درصد افزایش خواهد یافت. این افزایش دقت به دلیل همگرایی بهتر مدل‌ها و کاهش تفاوت‌های آماری بین داده‌های محلی است. به عبارتی دیگر، این روش کمک می‌کند که مدل‌ها با یکدیگر هماهنگ‌تر شوند و نتایج دقیق‌تری ارائه دهند.

با این حال، باید توجه داشت که اشتراک‌گذاری داده‌ها بین کاربران می‌تواند مسائل حریم شخصی را به



شکل ۳-۱: نمایش نحوه به اشتراک گذاری داده [۱۸].

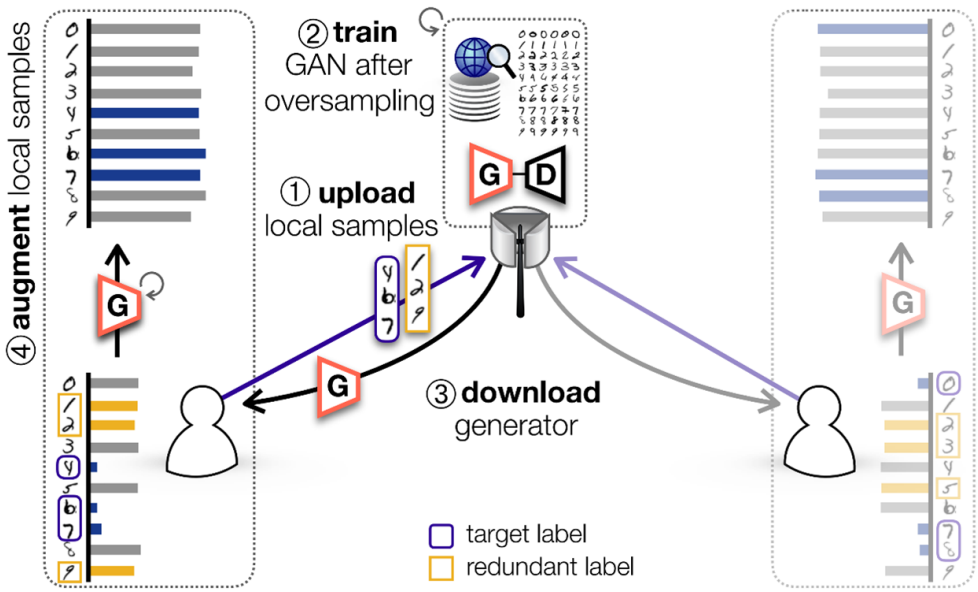
همراه داشته باشد. به عبارت دیگر، هنگامی که داده‌های اشتراکی بین کاربران مبادله می‌شود، احتمال نقض حریم شخصی کاربران افزایش می‌یابد. بنابراین، هنگام پیاده‌سازی این روش، ضروری است که اقدامات لازم برای حفظ حریم شخصی کاربران به طور جدی مد نظر قرار گیرد. این اقدامات می‌تواند شامل استفاده از تکنیک‌های رمزنگاری، ناشناس‌سازی داده‌ها، یا روش‌های دیگر برای محافظت از اطلاعات حساس کاربران باشد [۲۰].

در نهایت، روش به اشتراک‌گذاری داده‌ها بین کاربران، اگرچه می‌تواند به بهبود دقت و کارایی مدل‌ها کمک کند، اما نیازمند دقت و توجه ویژه‌ای به مسائل حریم شخصی است. پژوهشگران و توسعه‌دهندگان باید با در نظر گرفتن این چالش‌ها، راهکارهایی را برای حفظ امنیت و حریم شخصی کاربران در هنگام اجرای این روش‌ها ارائه دهند.

۳-۲-۲ بهبود داده

ابتدا، کاربران تعدادی از داده‌های خود را به سمت سرور ارسال می‌کنند. سرور، با استفاده از داده‌های دریافتی، یک مدل شبکه مولد رقابتی[1] ایجاد می‌کند و این مدل را برای تمامی کاربران ارسال می‌نماید. کاربران با استفاده از این شبکه مولد رقابتی و با توجه به داده‌های خود، تعدادی داده جدید تولید کرده و در مراحل بعدی آموزش از این داده‌ها نیز استفاده می‌کنند. به این ترتیب، شبکه مولد رقابتی به کاربران کمک می‌کند تا داده‌های بیشتری برای آموزش مدل‌های خود در اختیار داشته باشند و از این داده‌ها برای بهبود عملکرد مدل‌های خود استفاده کنند.

[1] Generative Adversarial Network (GAN)



شکل ۳-۲: استفاده از شبکه مولد رقابتی جهت تولید داده [۲۱].

در شکل ۳-۲ نحوه عملکرد این روش به تصویر کشیده شده است. این روش، به دلیل استفاده از تکنیک‌های رمزگذاری [۱] و رمزگشایی [۲] داده‌ها، نسبت به روش‌های اشتراک‌گذاری داده‌ها از نظر حفظ حریم شخصی کاربران بهتر عمل می‌کند. به این معنی که، به جای ارسال داده‌های خام کاربران به سرور یا دیگر کاربران، از داده‌های تولید شده توسط شبکه مولد رقابتی استفاده می‌شود که احتمال نقض حریم شخصی را کاهش می‌دهد.

استفاده از تکنیک‌های رمزگذاری و رمزگشایی داده‌ها در این روش، باعث می‌شود که داده‌های حساس کاربران در طول فرآیند آموزش، به صورت امن باقی بمانند. به عبارت دیگر، حتی اگر داده‌ها در طول انتقال یا در سرور مورد دسترسی غیرمجاز قرار گیرند، به دلیل رمزگذاری، اطلاعات واقعی کاربران فاش نخواهد شد. این ویژگی، امنیت و حریم شخصی کاربران را به طور قابل توجهی افزایش می‌دهد و از اطلاعات حساس آنان در برابر تهدیدات محافظت می‌کند [۲۱].

بنابراین، روش‌های بهبود داده شده که مبتنی بر رمزگذاری و رمزگشایی داده‌ها هستند، نه تنها به بهبود عملکرد مدل‌های یادگیری کمک می‌کنند، بلکه از حریم شخصی کاربران نیز حفاظت می‌نمایند. این ترکیب از امنیت و کارایی، این روش‌ها را به گزینه‌های مناسبی برای استفاده در سیستم‌های یادگیری فدرال تبدیل کرده است.

۳-۲-۳ تست

تست

[1] Encoding
[2] Decoding

۳-۲-۴ تست

تست

۳-۳ نگرش برپایه مدل

تست

۳-۴ نگرش برپایه چهارچوب

تست

۳-۵ نگرش برپایه الگوریتم

تست

مراجع

[1] Elbir, Ahmet M, Coleri, Sinem, Papazafeiropoulos, Anastasios K, Kourtessis, Pandelis, and Chatzinotas, Symeon. A family of hybrid federated and centralizedlearning architectures in machine learning. *IEEE Transactions on Cognitive Communications and Networking*, 2022.

[2] Zhou, Zhi, Chen, Xu, Li, En, Zeng, Liekang, Luo, Ke, and Zhang, Junshan. Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings of the IEEE*, 107(8):1738–1762, 2019.

[3] Tomas. Decentralized x: Aggregating heterogeneous and decentralized ais. https://www.omron.com/global/en/technology/information/dcx. [Accessed: 17 Apr 2024].

[4] Smith, Virginia, Chiang, Chao-Kai, Sanjabi, Maziar, and Talwalkar, Ameet S. Federated multi-task learning. *Advances in neural information processing systems*, 30, 2017.

[5] McMahan, Brendan, Ramage Daniel. Federated learning: Collaborative machine learning without centralized training data. https://www.omron.com/global/en/technology/information/dcx, 6 Apr 2017. [Accessed: 18 Apr 2024].

[6] Li, Tian, Sahu, Anit Kumar, Talwalkar, Ameet, and Smith, Virginia. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3):50–60, 2020.

[7] Talaei, Mahtab. Algorithm development and performance analysis for adaptive differential privacy in federated learning, 21 Aug 2022.

[8] Kim, Jiyeon, Yang, Inseok, and Lee, Dongik. Control allocation based compensation for faulty blade actuator of wind turbine. *IFAC Proceedings Volumes*, 45(20):355–360, 2012.

[9] Wang, Hongyi, Sievert, Scott, Liu, Shengchao, Charles, Zachary, Papailiopoulos, Dimitris, and Wright, Stephen. Atomo: Communication-efficient learning via atomic sparsification. *Advances in neural information processing systems*, 31, 2018.

[10] Konečnỳ, Jakub, McMahan, H Brendan, Yu, Felix X, Richtárik, Peter, Suresh, Ananda Theertha, and Bacon, Dave. Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*, 2016.

[11] Fang, Chen, Guo, Yuanbo, Hu, Yongjin, Ma, Bowen, Feng, Li, and Yin, Anqi. Privacy-preserving and communication-efficient federated learning in internet of things. *Computers & Security*, 103:102199, 2021.

[12] Konečnỳ, Jakub, McMahan, Brendan, and Ramage, Daniel. Federated optimization: Distributed optimization beyond the datacenter. *arXiv preprint arXiv:1511.03575*, 2015.

[13] Hasan, Jahid. Security and privacy issues of federated learning. *arXiv preprint arXiv:2307.12181*, 2023.

[14] Yin, Xuefei, Zhu, Yanming, and Hu, Jiankun. A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions. *ACM Computing Surveys (CSUR)*, 54(6):1–36, 2021.

[15] McMahan, Brendan, Moore, Eider, Ramage, Daniel, Hampson, Seth, and y Arcas, Blaise Aguera. Communication-efficient learning of deep networks from decentralized data. in *Artificial intelligence and statistics*, pp. 1273–1282. PMLR, 2017.

[16] Ioffe, Sergey and Szegedy, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. in *International conference on machine learning*, pp. 448–456. pmlr, 2015.

[17] Li, Tian, Sahu, Anit Kumar, Zaheer, Manzil, Sanjabi, Maziar, Talwalkar, Ameet, and Smith, Virginia. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems*, 2:429–450, 2020.

[18] Zhao, Yue, Li, Meng, Lai, Liangzhen, Suda, Naveen, Civin, Damon, and Chandra, Vikas. Federated learning with non-iid data. *arXiv preprint arXiv:1806.00582*, 2018.

[19] Collins, Liam, Hassani, Hamed, Mokhtari, Aryan, and Shakkottai, Sanjay. Exploiting shared representations for personalized federated learning. in *International conference on machine learning*, pp. 2089–2099. PMLR, 2021.

[20] Ma, Xiaodong, Zhu, Jia, Lin, Zhihao, Chen, Shanxuan, and Qin, Yangjie. A state-of-the-art survey on solving non-iid data in federated learning. *Future Generation Computer Systems*, 135:244–258, 2022.

[21] Jeong, Eunjeong, Oh, Seungeun, Kim, Hyesung, Park, Jihong, Bennis, Mehdi, and Kim, Seong-Lyun. Communication-efficient on-device machine learning: Federated distillation and augmentation under non-iid private data. *arXiv preprint arXiv:1811.11479*, 2018.