***"Scaling Language Model Size in Cross-Device Federated Learning"***

روش‌های جدیدی را برای آموزش مدل‌های بزرگ زبانی در چارچوب یادگیری فدرال ارائه می‌دهد. از آنجا که مدل‌های یادگیری فدرال (FL) به صورت توزیع‌شده و بر روی دستگاه‌های محدود مانند تلفن‌های همراه اجرا می‌شوند، بهینه‌سازی‌هایی برای کاهش هزینه‌های محاسباتی و ارتباطی لازم است. این مقاله مدل‌های ترنسفورمر و کانفرمر بزرگ را با استفاده از تکنیک‌های پیشرفته‌ای مانند **کوانتیزاسیون، آموزش جزئی مدل، یادگیری انتقالی، و بهینه‌سازی‌های کارآمد** بهینه‌سازی می‌کند و توانایی اجرای مدل‌هایی با دقت بالا و ارتباطات کمتر را در محیط یادگیری فدرال نشان می‌دهد.

**ساختار و محتوای اصلی مقاله:**

1. **مقدمه**: بیان محدودیت‌های کنونی در یادگیری فدرال، به‌ویژه در آموزش مدل‌های بزرگ زبانی بر روی دستگاه‌های با منابع محدود.
2. **داده‌ها و مدل‌ها**: استفاده از داده‌های **Stack Overflow** و بررسی چندین معماری مدل (LSTM، Transformer، و Conformer) برای آموزش.
3. **کاهش هزینه در هر دور ارتباطی**: استفاده از کوانتیزاسیون و آموزش جزئی مدل برای کاهش هزینه محاسبات و ارتباطات بین سرور و دستگاه‌های کاربران.
4. **کاهش تعداد دورهای ارتباطی**: استفاده از یادگیری انتقالی (pretraining) بر روی مجموعه داده‌های متنی بزرگ برای کاهش تعداد دورهای لازم تا همگرایی مدل.
5. **ترکیب تکنیک‌ها**: بهینه‌سازی‌های مختلف مانند کوانتیزاسیون و یادگیری انتقالی برای آموزش مدل‌های بزرگ‌تر و کارآمدتر.

**نتیجه‌گیری:**

این مطالعه نشان می‌دهد که تکنیک‌های مختلف می‌توانند به بهبود اجرای مدل‌های بزرگ در یادگیری فدرال کمک کنند، و امکان اجرای مدل‌های ترنسفورمر بزرگ‌تر را با هزینه ارتباطی و محاسباتی کمتر فراهم کنند.

چکیده این مقاله به بررسی راهکارهایی برای آموزش مدل‌های زبانی بزرگ در چارچوب یادگیری فدرال در دستگاه‌های محدود مانند تلفن‌های همراه می‌پردازد. با استفاده از تکنیک‌هایی مانند **آموزش جزئی مدل، کوانتیزاسیون، یادگیری انتقالی، و بهینه‌سازی ارتباطات**، این تحقیق نشان می‌دهد که می‌توان مدل‌هایی مانند ترنسفورمر و کانفرمر را با حدود ۲۵ میلیون پارامتر به‌طور کارآمد آموزش داد. این مدل‌ها دقت بالاتری نسبت به مدل‌های کوچک‌تر دارند و نیازهای محاسباتی و ارتباطی بسیار کمتری دارند، در نتیجه کارآمدی آن‌ها در یادگیری فدرال افزایش می‌یابد.

این مقاله به بررسی چالش‌های یادگیری فدرال در محیط‌های محدود مانند دستگاه‌های کاربران (مثل تلفن‌های همراه) می‌پردازد. یادگیری فدرال روشی است که به دستگاه‌ها اجازه می‌دهد بدون انتقال داده‌های کاربران به سرور مرکزی، در آموزش مدل مشارکت کنند و بدین ترتیب حریم خصوصی حفظ می‌شود. این روش برای کاربردهایی مثل کیبوردهای مجازی و تشخیص گفتار استفاده می‌شود.

یکی از محدودیت‌های اصلی در یادگیری فدرال، اجرای مدل‌های بزرگ زبانی بر روی دستگاه‌های لبه‌ای است، زیرا این دستگاه‌ها توان محاسباتی و ارتباطی محدودی دارند. در نتیجه، مدل‌هایی که به‌طور معمول استفاده می‌شوند کوچک‌تر از مدل‌های سرور-محور هستند و شامل مدل‌هایی با کمتر از ۴ میلیون پارامتر مانند LSTM می‌باشند. با توجه به محدودیت‌های دستگاه‌ها، این مقاله به معرفی تکنیک‌هایی پرداخته که به کاهش هزینه‌های محاسباتی و ارتباطی کمک می‌کنند و امکان آموزش مدل‌های بزرگ‌تر مانند ترنسفورمر و کانفرمر را فراهم می‌سازند.

روش‌های استفاده‌شده در این مقاله شامل کوانتیزاسیون (فشرده‌سازی پارامترها)، آموزش جزئی (آموزش بخشی از مدل به‌جای کل آن)، و یادگیری انتقالی (انتقال اطلاعات از مدل‌های پیش‌آموزش‌دیده) می‌باشد. این تکنیک‌ها به مدل اجازه می‌دهند تا با کاهش نیاز به ارتباطات بین سرور و دستگاه، بتواند دقت بالاتری را به دست آورد.

در بخش "مجموعه داده‌ها و مدل‌ها"، مقاله به تشریح داده‌ها و مدل‌های مورد استفاده برای آزمایش می‌پردازد. مجموعه داده‌ها از **Stack Overflow** انتخاب شده‌اند، جایی که پست‌های کاربران به‌صورت فدرال و براساس نام کاربری گروه‌بندی شده‌اند. برای پیش‌پردازش متون، از مدل **sentence-piece** برای شکستن کلمات به واحدهای کوچکتر با اندازه واژگان ۴۰۰۰ استفاده شده است.

سه نوع مدل اصلی در این تحقیق مورد آزمایش قرار گرفته‌اند:

1. **LSTM** (مدل حافظه طولانی-کوتاه)
2. **Transformer**
3. **Conformer**

مدل‌های کوچک و بزرگ از هر کدام از این معماری‌ها با حداکثر ۲۵ میلیون پارامتر ساخته شده‌اند، از جمله LSTM با ۱۸.۸ میلیون، Transformer با ۲۱ میلیون، و Conformer با ۲۰.۲ میلیون پارامتر.

برای ارزیابی عملکرد مدل‌ها، **پریشانی آزمون (test perplexity)** به‌عنوان معیار اصلی استفاده شده است و تنظیمات بهینه برای هر مدل و اندازه آن براساس این معیار انتخاب شده‌اند. آزمایش‌ها با استفاده از کتابخانه‌های JAX و FedJAX برای شبیه‌سازی یادگیری فدرال پیاده‌سازی شده‌اند.

در بخش چهارم مقاله با عنوان **"هزینه هر دور"**، به بررسی تکنیک‌هایی برای کاهش هزینه‌های محاسباتی و ارتباطی مدل‌های بزرگ در هر دور ارتباطی پرداخته شده است. به دلیل محدودیت منابع در دستگاه‌های کاربران، بهینه‌سازی‌هایی برای کاهش بار ارتباطی و محاسباتی ضروری است. دو روش اصلی برای کاهش این هزینه‌ها عبارتند از:

1. **آموزش جزئی مدل (Partial Model Training)**: به جای آموزش کل مدل، تنها بخشی از آن آموزش داده می‌شود. این رویکرد باعث کاهش پارامترهایی می‌شود که باید در هر دور انتقال داده شوند، و همچنین هزینه محاسباتی روی دستگاه‌ها را کاهش می‌دهد. در این مقاله، آموزش جزئی مدل به روش "Partial Variable Training" انجام شده که درصد خاصی از پارامترهای مدل را در هر دور ثابت نگه می‌دارد و تنها بخش‌هایی از مدل را آموزش می‌دهد.
2. **کوانتیزاسیون (Quantization)**: این روش با کاهش تعداد بیت‌های مورد نیاز برای نمایش پارامترها، حجم داده‌های انتقالی بین سرور و دستگاه‌ها را کاهش می‌دهد. در این مقاله از کوانتیزاسیون تصادفی با سطح k ثابت استفاده شده است که پارامترها را به نزدیک‌ترین سطح مقدار تنظیم می‌کند. کوانتیزاسیون در اینجا برای کاهش هزینه انتقال داده از سرور به دستگاه (دانلود) و از دستگاه به سرور (آپلود) استفاده می‌شود.

**نتایج:**

این روش‌ها به میزان قابل توجهی هزینه ارتباطی و محاسباتی هر دور را کاهش دادند. به طور خاص، استفاده از کوانتیزاسیون باعث کاهش ۶۲.۵٪ در هزینه ارتباطی کل و آموزش جزئی مدل، باعث کاهش ۷۰٪ در هزینه‌های آپلود شد.

در بخش پنجم، با عنوان **"تعداد دورهای ارتباطی"**، مقاله به تکنیک‌هایی برای کاهش تعداد دورهای ارتباطی مورد نیاز در آموزش مدل می‌پردازد. این کاهش باعث می‌شود که مدل سریع‌تر به همگرایی برسد و هزینه‌های ارتباطی کاهش یابد. دو روش اصلی که در این بخش بررسی شده‌اند عبارتند از:

1. **یادگیری انتقالی (Transfer Learning)**: با پیش‌آموزش مدل بر روی یک مجموعه داده مشابه قبل از آموزش فدرال، تعداد دورهای مورد نیاز برای رسیدن به دقت مطلوب کاهش می‌یابد. در این تحقیق، از مجموعه‌ داده‌های بزرگ کتاب‌های دیجیتال و مجموعه داده LM1B برای پیش‌آموزش استفاده شده است. نتایج نشان داد که پیش‌آموزش بر روی مجموعه کتاب‌ها عملکرد بهتری نسبت به LM1B داشت و تعداد دورهای ارتباطی لازم را تا ۳ برابر کاهش داد.
2. **استفاده از بهینه‌سازهای مختلف**: مقاله همچنین الگوریتم‌های بهینه‌سازی خاص فدرال، مانند **FedProx** و **MimeLite** را بررسی کرده است. هدف این بهینه‌سازها کاهش تعداد دورهای ارتباطی تا رسیدن به همگرایی مدل بود. با این حال، نتایج نشان داد که این الگوریتم‌ها در مقایسه با الگوریتم پایه **FedAvg** تفاوت چشمگیری در سرعت همگرایی ایجاد نمی‌کنند.

### نتیجه

این روش‌ها نشان می‌دهند که با استفاده از یادگیری انتقالی، می‌توان تعداد دورهای ارتباطی را کاهش داد و کارایی مدل‌ها را بهبود بخشید، در حالی که الگوریتم‌های بهینه‌سازی فدرال تفاوت چشمگیری در کاهش دورهای ارتباطی ایجاد نکردند.

**"Scaling Federated Learning for Fine-tuning of Large Language Models"**

به بررسی چالش‌ها و راهکارهای یادگیری فدرال (FL) در آموزش و تنظیم مدل‌های بزرگ زبانی مانند BERT، ALBERT و DistilBERT می‌پردازد. در این مقاله، مدل‌ها بر روی سه مجموعه داده مختلف (IMDB، Yelp F، و AG News) آزمایش شده‌اند. مقاله شامل بخش‌های زیر است:

**1. چکیده:**

چکیده به این موضوع اشاره دارد که یادگیری فدرال می‌تواند به حفظ حریم خصوصی کمک کند و امکان به‌کارگیری مدل‌های زبانی بزرگ را در محیط‌های توزیع‌شده فراهم می‌کند. مقاله عملکرد سه مدل BERT-variant را در تنظیمات مختلف یادگیری فدرال بررسی کرده و تأثیر تعداد مشتریان بر عملکرد مدل را ارزیابی می‌کند. یافته‌ها نشان می‌دهد که اندازه بزرگ مدل‌ها مانعی برای یادگیری فدرال نیست، اما مدل‌ها رفتار متفاوتی در برابر روش **Federated Averaging (FedAvg)** نشان می‌دهند. به‌ویژه، DistilBERT کندتر همگرا می‌شود و حتی ممکن است عملکرد تصادفی ارائه دهد.

**2. مقدمه:**

در مقدمه، مقاله به معرفی معماری‌های ترنسفورمر و پیشرفت‌های آن‌ها در وظایف زبانی مختلف می‌پردازد. مدل‌های مبتنی بر ترنسفورمر مانند BERT، ALBERT و DistilBERT در کارهایی مانند تحلیل احساسات، طبقه‌بندی اسناد و پرسش و پاسخ به نتایج عالی دست یافته‌اند. در این مقاله، هدف بررسی تنظیم دقیق (Fine-tuning) مدل‌های ترنسفورمر در چارچوب یادگیری فدرال است. یادگیری فدرال به آموزش غیرمتمرکز مدل‌ها اشاره دارد که در آن مدل‌ها به جای انتقال داده‌ها، روی دستگاه‌های محلی کاربران آموزش می‌بینند. این روش به کاهش نگرانی‌های مرتبط با حفظ حریم خصوصی کمک می‌کند.

**3. کارهای مرتبط:**

در این بخش، مقاله به تحقیقات پیشین در زمینه یادگیری فدرال و بهینه‌سازی مدل‌ها اشاره می‌کند. روش FedAvg (McMahan et al., 2017) به‌عنوان روشی برای ترکیب به‌روزرسانی‌های مدل از مشتریان مختلف معرفی می‌شود. مطالعات پیشین به بهبود کارایی یادگیری فدرال در داده‌های ناهمگون و حفظ حریم خصوصی کاربران پرداخته‌اند. مقالات اخیر نیز به بررسی چگونگی تنظیم دقیق مدل‌های ترنسفورمر در یادگیری فدرال اشاره می‌کنند، اما تعداد کمی از آن‌ها به بررسی مدل‌های بزرگ در این تنظیمات پرداخته‌اند.

**4. روش‌ها:**

در این بخش، روش یادگیری فدرال برای آموزش مدل‌های ترنسفورمر توضیح داده شده است. از روش **FedAvg** برای ترکیب به‌روزرسانی‌های مشتریان استفاده می‌شود. سه مدل زبانی مختلف شامل BERT (110 میلیون پارامتر)، ALBERT (11 میلیون پارامتر) و DistilBERT (65 میلیون پارامتر) انتخاب شده‌اند. این مدل‌ها از کتابخانه Huggingface گرفته شده و با استفاده از PyTorch پیاده‌سازی شده‌اند. مدل‌ها ابتدا پیش‌آموزش شده و سپس در محیط یادگیری فدرال تنظیم دقیق می‌شوند.

**مجموعه داده‌ها** شامل:

* **IMDB**: داده‌های بررسی فیلم با ۵۰٬۰۰۰ نمونه.
* **Yelp F**: بررسی کسب‌وکارهای محلی با رتبه‌بندی ۱ تا ۵.
* **AG News**: مقالات خبری با ۱۲۰٬۰۰۰ نمونه.

**5. نتایج:**

مقاله نشان می‌دهد که هرچه تعداد مشتریان بیشتر باشد، عملکرد مدل‌ها کاهش می‌یابد. BERT و ALBERT تا ۳۲ مشتری مقیاس می‌شوند و کاهش شدیدی در عملکرد ندارند، اما DistilBERT پس از ۴ مشتری به عملکرد تصادفی می‌رسد. این نشان می‌دهد که مدل‌های کوچکتر مانند DistilBERT حساسیت بیشتری به افزایش تعداد مشتریان دارند.

**6. بحث و نتیجه‌گیری:**

این مقاله نشان می‌دهد که یادگیری فدرال می‌تواند برای تنظیم دقیق مدل‌های بزرگ زبانی استفاده شود و مدل‌هایی مانند BERT و ALBERT عملکرد خوبی حتی در تنظیمات توزیع‌شده ارائه می‌دهند. اما DistilBERT به دلیل معماری متفاوت و فرآیند فشرده‌سازی در مرحله پیش‌آموزش، به تنظیمات یادگیری فدرال حساس‌تر است. این مقاله همچنین به مشکلاتی مانند ناپایداری و نیاز به بهبود در روش‌های ترکیب مدل‌ها برای بهبود عملکرد در تنظیمات فدرال اشاره می‌کند.

**نتیجه‌گیری:**

این مقاله به بررسی تنظیم دقیق مدل‌های بزرگ زبانی در چارچوب یادگیری فدرال پرداخته و نشان داده است که یادگیری فدرال برای این مدل‌ها مناسب است، اما چالش‌هایی نیز وجود دارد که نیازمند تحقیقات بیشتر است.

**"Low-Parameter Federated Learning with Large Language Models"**

به بررسی یادگیری فدرال با استفاده از مدل‌های زبانی بزرگ پرداخته و روشی به نام **Low-Parameter Federated Learning (LP-FL)** را معرفی می‌کند که برای کاهش هزینه‌های محاسباتی و ارتباطی طراحی شده است. من در ادامه بخش‌های مختلف مقاله را به تفصیل توضیح می‌دهم:

### 1. چکیده:

این مقاله یادگیری فدرال را برای وظایف **درک زبان طبیعی (NLU)** با استفاده از مدل‌های زبانی بزرگ (LLMs) بررسی می‌کند. یادگیری فدرال به دلیل منابع محدود دستگاه‌های موبایل، بهینه‌سازی‌های بیشتری نیاز دارد. مقاله به چالش‌های تنظیم دقیق (fine-tuning) مدل‌های زبانی بزرگ با استفاده از داده‌های اندک اشاره می‌کند و روش **LP-FL** را معرفی می‌کند که شامل **یادگیری با استفاده از داده‌های دارای برچسب نرم (soft labels)** و **تکنیک Low-Rank Adaptation (LoRA)** است. روش پیشنهادی هزینه‌های محاسباتی را کاهش داده و عملکرد بهتری نسبت به تنظیم دقیق با پارامترهای کامل (FP-FL) نشان می‌دهد.

### 2. مقدمه:

در مقدمه، مقاله به رشد مدل‌های زبانی بزرگ مانند **GPT-3** و **BERT** و تأثیر آن‌ها بر هوش مصنوعی اشاره می‌کند. این مدل‌ها با وجود دقت بالای خود، به منابع محاسباتی زیادی نیاز دارند که اجرای آن‌ها بر روی دستگاه‌های لبه‌ای مانند تلفن‌های همراه دشوار است. یادگیری فدرال، که داده‌ها را بر روی دستگاه‌های محلی پردازش می‌کند، می‌تواند به حفظ حریم خصوصی کمک کند. اما به دلیل اندازه بزرگ مدل‌ها و منابع محدود دستگاه‌ها، تنظیم دقیق این مدل‌ها در یادگیری فدرال چالش‌برانگیز است. مقاله به بررسی روشی برای کاهش پارامترهای قابل تنظیم و بهبود کارایی یادگیری فدرال می‌پردازد.

### 3. کارهای مرتبط:

در این بخش، مقاله به تحقیقات پیشین در زمینه یادگیری فدرال و تنظیم دقیق مدل‌های زبانی اشاره می‌کند. روش‌هایی مانند **FedAvg**، **FedFSL** و **FedSSL** معرفی شده‌اند که برای یادگیری فدرال با داده‌های محدود و برچسب‌گذاری نشده استفاده می‌شوند. همچنین به تکنیک‌های تنظیم دقیق پارامترها (PEFT) مانند **Prefix-Tuning** و **LoRA** اشاره شده است که بدون تنظیم همه پارامترهای مدل، عملکردی مشابه تنظیم دقیق کامل ارائه می‌دهند. این مقاله از تکنیک LoRA برای کاهش پارامترهای قابل تنظیم در یادگیری فدرال استفاده می‌کند.

در این بخش، مقاله به بررسی تحقیقات قبلی در زمینه **یادگیری فدرال** و **تنظیم دقیق مدل‌های زبانی بزرگ** می‌پردازد. یادگیری فدرال به‌عنوان یک روش یادگیری توزیع‌شده معرفی می‌شود که مدل‌ها را بر روی داده‌های غیرمتمرکز آموزش می‌دهد. روش‌های مختلفی مانند **FedAvg**، **FedFSL** و **FedSSL** معرفی شده‌اند که به بهبود کارایی مدل‌ها در محیط‌های با داده‌های محدود و برچسب‌گذاری نشده کمک می‌کنند.

در مورد مدل‌های زبانی بزرگ (LLMs) نیز به تحقیقات پیشین مانند **GPT-3** و **BERT** اشاره شده است که قدرت فوق‌العاده‌ای در وظایف زبانی دارند. با این حال، این مدل‌ها نیازمند منابع محاسباتی بالایی هستند. تکنیک‌های تنظیم دقیق پارامترها (PEFT) مانند **Prefix-Tuning** و **LoRA** معرفی شده‌اند که به جای تنظیم کامل پارامترهای مدل، تنها بخشی از پارامترها را تنظیم می‌کنند و هزینه محاسباتی را کاهش می‌دهند.

### 4. رویکرد:

روش **Low-Parameter Federated Learning (LP-FL)** به جای تنظیم کامل پارامترها، تنها بخشی از پارامترهای مدل را تنظیم می‌کند. در این روش از **LoRA** استفاده می‌شود که یک لایه اضافی به مدل اضافه می‌کند و تنها پارامترهای این لایه آموزش داده می‌شوند. در هر دور از یادگیری، مشتریان (دستگاه‌ها) مدل را با داده‌های محلی خود تنظیم می‌کنند و پارامترهای LoRA را به سرور ارسال می‌کنند. سرور از روش **FedAvg** برای تجمیع این پارامترها استفاده کرده و آن‌ها را به مشتریان بازمی‌گرداند. همچنین، داده‌های بدون برچسب با استفاده از برچسب‌های نرم (soft labels) برچسب‌گذاری می‌شوند تا مجموعه داده‌های برچسب‌دار گسترش یابد.

مقاله یک روش جدید به نام **Low-Parameter Federated Learning (LP-FL)** را پیشنهاد می‌کند. این روش بر اساس تکنیک **Low-Rank Adaptation (LoRA)** است که تنها بخشی از پارامترهای مدل را تنظیم می‌کند و پارامترهای اصلی را ثابت نگه می‌دارد. این رویکرد شامل مراحل زیر است:

1. **تنظیم دقیق پارامترها**: به جای تنظیم کامل مدل، تنها پارامترهای LoRA تنظیم می‌شوند. این باعث کاهش بار محاسباتی و ارتباطی می‌شود.
2. **استفاده از برچسب‌های نرم (Soft Labels)**: داده‌های بدون برچسب با استفاده از مدل‌های پیش‌آموزش‌دیده برچسب‌گذاری می‌شوند. این برچسب‌های نرم به افزایش مجموعه داده‌های برچسب‌دار کمک می‌کنند.
3. **تجمیع پارامترها**: پارامترهای تنظیم‌شده از تمامی مشتریان (دستگاه‌ها) به سرور ارسال می‌شوند و از روش **FedAvg** برای ترکیب آن‌ها استفاده می‌شود. سپس پارامترهای به‌روزشده به مشتریان بازگردانده می‌شوند.

این رویکرد به دلیل استفاده از تعداد کمتری پارامتر، هزینه‌های محاسباتی را کاهش داده و از **بیش‌برازش (overfitting)** جلوگیری می‌کند. همچنین، استفاده از برچسب‌های نرم به بهبود کیفیت داده‌ها و افزایش دقت مدل کمک می‌کند.

### 5. آزمایش‌ها:

آزمایش‌ها بر روی دو مجموعه داده **IMDB** و **Yelp** برای طبقه‌بندی احساسات انجام شده است. داده‌های این مجموعه‌ها به دو دسته **برچسب‌دار (labeled)** و **بدون برچسب (unlabeled)** تقسیم شده‌اند. مدل مورد استفاده **BERT-Large** با 336 میلیون پارامتر است. آزمایش‌ها نشان می‌دهند که روش LP-FL با تنظیم تنها ۳۰٪ از پارامترها، عملکردی مشابه یا حتی بهتر از روش تنظیم کامل پارامترها (FP-FL) دارد، اما با هزینه محاسباتی و ارتباطی بسیار کمتر. نتایج نشان داد که LP-FL به طور متوسط تنها ۱ تا ۲ درصد کمتر از تنظیم کامل پارامترها عمل می‌کند.

### 6. نتیجه‌گیری:

مقاله نتیجه می‌گیرد که روش **LP-FL** می‌تواند به طور مؤثری تنظیم دقیق مدل‌های زبانی بزرگ را در یادگیری فدرال انجام دهد، بدون نیاز به تنظیم کامل پارامترها. این روش می‌تواند بر محدودیت‌های دستگاه‌های لبه‌ای غلبه کند و حتی در تنظیمات با داده‌های محدود نیز عملکرد خوبی ارائه دهد. مقاله به این نتیجه می‌رسد که **یادگیری با پارامترهای کم** می‌تواند جایگزین مناسبی برای تنظیم دقیق کامل در محیط‌های یادگیری فدرال باشد و کاربردهای امیدوارکننده‌ای در دستگاه‌های موبایل داشته باشد.

**«OpenFedLLM: Training Large Language Models on Decentralized Private Data via Federated Learning» ا**

### چکیده

در چکیده، نویسندگان به معرفی OpenFedLLM می‌پردازند که یک فریم‌ورک برای آموزش مدل‌های زبانی بزرگ (LLMs) به‌صورت مشترک و با حفظ حریم خصوصی با استفاده از یادگیری فدرال (FL) است. این مقاله به چالش‌های کمبود داده‌های عمومی با کیفیت بالا اشاره می‌کند و بر اهمیت داده‌های خصوصی توزیع‌شده که نمی‌توانند به‌صورت عمومی به اشتراک گذاشته شوند، تأکید می‌کند. OpenFedLLM شامل الگوریتم‌های مختلف FL و روش‌های تنظیم دستورالعمل و هم‌راستایی ارزش‌ها است. نتایج تجربی نشان‌دهنده بهبود قابل توجه عملکرد مدل‌های آموزشی از طریق FL نسبت به آموزش محلی است. به‌ویژه، مدل‌های آموزش‌دیده با FL می‌توانند در برخی حوزه‌ها مانند مالی، عملکرد بهتری نسبت به مدل‌های شناخته‌شده مانند GPT-4 داشته باشند.

### مقدمه

در مقدمه، نویسندگان به موفقیت مدل‌های زبانی بزرگ در زمینه‌های مختلف اشاره می‌کنند و بر چالش‌های مربوط به کمبود داده‌های عمومی با کیفیت تأکید می‌کنند. آنها بیان می‌کنند که پیش‌بینی می‌شود داده‌های عمومی با کیفیت تا سال 2026 به پایان برسد و این مسئله می‌تواند توسعه LLMها را به چالش بکشد. نویسندگان بر اهمیت همکاری در آموزش LLMها با استفاده از داده‌های خصوصی توزیع‌شده تأکید می‌کنند، به‌طوری که این داده‌ها به‌دلیل مسائل حریم خصوصی نمی‌توانند به‌طور عمومی به اشتراک گذاشته شوند. در نهایت، آنها به استفاده از یادگیری فدرال به‌عنوان یک راه‌حل برای این چالش پرداخته و به تشریح چهار مرحله اصلی در این فرایند می‌پردازند.

### کارهای مرتبط

در بخش کارهای مرتبط، نویسندگان به مرور ادبیات موجود در زمینه LLMها و یادگیری فدرال می‌پردازند. آنها به بررسی مراحل آموزش مدل‌های زبانی بزرگ، از جمله پیش‌آموزش، تنظیم دستورالعمل و هم‌راستایی ارزش‌ها می‌پردازند. همچنین، نویسندگان به مشکلات مرتبط با داده‌های عمومی و تلاش‌های اخیر برای استفاده از داده‌های خصوصی اشاره می‌کنند. در زمینه یادگیری فدرال، آن‌ها به الگوریتم‌های مختلف و چالش‌های مرتبط با داده‌های ناهمگن اشاره می‌کنند و به این نکته می‌پردازند که اکثر تحقیقات قبلی در زمینه کاربرد FL بر روی مدل‌های کوچک‌تر و وظایف ساده‌تر بوده‌اند.

### چهارچوب OpenFedLLM

در این بخش، نویسندگان به تشریح چهارچوب OpenFedLLM می‌پردازند. این چهارچوب شامل مراحل مختلف آموزش مدل‌های زبانی بزرگ از طریق FL است. آنها توضیح می‌دهند که چگونه OpenFedLLM به استانداردهای FL مانند تجمیع امن و حریم خصوصی پیروی می‌کند. فرایند آموزش شامل چهار مرحله کلیدی است: ارسال مدل جهانی از سرور به مشتریان، آموزش محلی مدل در هر مشتری، بارگذاری مدل محلی به سرور و تجمیع مدل‌های محلی برای به‌روزرسانی مدل جهانی. همچنین، روش‌های تنظیم دستورالعمل و هم‌راستایی ارزش‌ها به‌عنوان دو کاربرد کلیدی در این چهارچوب معرفی می‌شوند

**الگوریتم‌های FL در OpenFedLLM**

OpenFedLLM شامل هفت الگوریتم نماینده FL است که هر کدام به منظور بهبود عملکرد مدل‌های زبانی بزرگ در شرایط مختلف طراحی شده‌اند. این الگوریتم‌ها به شرح زیر هستند:

1. **FedAvg (Federated Averaging)**:
   * این الگوریتم اصلی‌ترین و شناخته‌شده‌ترین روش FL است. در FedAvg، مدل جهانی از سرور به همه مشتریان ارسال می‌شود و هر مشتری با استفاده از داده‌های محلی خود، مدل را آموزش می‌دهد. سپس، مدل‌های محلی به سرور ارسال می‌شوند و مدل جهانی با میانگین‌گیری از وزن‌ها به‌روزرسانی می‌شود.
   * **مزایا**: سادگی و کارایی در سناریوهای داده همگن.
   * **معایب**: عملکرد متوسط در شرایطی با داده‌های ناهمگن.
2. **FedProx**:
   * FedProx یک نسخه بهبود یافته از FedAvg است که به‌منظور مدیریت داده‌های ناهمگن طراحی شده است. این الگوریتم یک عبارت جریمه به تابع هزینه محلی اضافه می‌کند تا فاصله بین مدل محلی و مدل جهانی را کاهش دهد.
   * **مزایا**: قابلیت بهبود عملکرد در شرایط ناهمگن.
   * **معایب**: پیچیدگی محاسباتی بیشتر نسبت به FedAvg.
3. **SCAFFOLD**:
   * این الگوریتم به منظور اصلاح گرادیان‌های محلی برای بهبود دقت در یادگیری فدرال طراحی شده است. SCAFFOLD از متغیرهای اصلاحی برای کاهش اثرات داده‌های ناهمگن استفاده می‌کند.
   * **مزایا**: بهبود در دقت مدل جهانی با کاهش اثرات نوسانات محلی.
   * **معایب**: نیاز به پیاده‌سازی پیچیده‌تری دارد.
4. **FedAvgM (Federated Averaging with Momentum)**:
   * FedAvgM به الگوریتم FedAvg یک مؤلفه مومنتوم اضافه می‌کند که به بهبود پایداری یادگیری کمک می‌کند. این روش می‌تواند به تسریع همگرایی مدل جهانی کمک کند.
   * **مزایا**: کاهش نوسانات و بهبود همگرایی.
   * **معایب**: ممکن است در شرایط خاصی عملکرد بهتری نداشته باشد.
5. **FedAdagrad**:
   * این الگوریتم از تکنیک Adagrad برای به‌روزرسانی وزن‌ها استفاده می‌کند که به هر پارامتر یک نرخ یادگیری متغیر اختصاص می‌دهد. این روش به‌ویژه در شرایطی که داده‌ها ناهمگن هستند، می‌تواند مفید باشد.
   * **مزایا**: تطبیق‌پذیری با داده‌های ناهمگن.
   * **معایب**: ممکن است در برخی شرایط یادگیری ضعیف‌تری داشته باشد.
6. **FedYogi**:
   * FedYogi یک الگوریتم پیشرفته FL است که از تکنیک‌های بهینه‌سازی Yogi استفاده می‌کند. این الگوریتم به‌طور خاص برای مدیریت داده‌های ناهمگن طراحی شده است و می‌تواند سرعت یادگیری را افزایش دهد.
   * **مزایا**: بهبود سرعت یادگیری و دقت.
   * **معایب**: ممکن است پیچیدگی بیشتری در پیاده‌سازی داشته باشد.
7. **FedAdam**:
   * این الگوریتم بر اساس تکنیک Adam طراحی شده است و از مزایای آن در یادگیری فدرال استفاده می‌کند. FedAdam می‌تواند به خوبی با داده‌های ناهمگن سازگار شود و عملکرد بهتری نسبت به روش‌های سنتی ارائه دهد.
   * **مزایا**: عملکرد بهتر در سناریوهای پیچیده و ناهمگن.
   * **معایب**: نیاز به منابع محاسباتی بیشتر نسبت به روش‌های ساده‌تر.

**نتیجه‌گیری**

استفاده از این الگوریتم‌ها در OpenFedLLM به‌منظور بهبود عملکرد مدل‌های زبانی بزرگ در شرایط مختلف، به‌ویژه در مواجهه با داده‌های ناهمگن است. نتایج تجربی نشان می‌دهد که این الگوریتم‌ها می‌توانند به‌طور مداوم بهتر از آموزش محلی عمل کنند و به سازمان‌ها انگیزه می‌دهند تا در یادگیری فدرال مشارکت کنند.

در OpenFedLLM، مقایسه عملکرد الگوریتم‌های مختلف یادگیری فدرال (FL) به‌منظور ارزیابی کارایی و قابلیت‌های آن‌ها در آموزش مدل‌های زبانی بزرگ انجام شده است. این مقایسه شامل تجزیه و تحلیل نتایج تجربی و ارزیابی‌های مختلف است. در زیر به برخی از نکات کلیدی در مورد مقایسه عملکرد این الگوریتم‌ها پرداخته می‌شود:

**1. عملکرد کلی در آموزش LLMها**

تجربیات نشان داده‌اند که به‌طور کلی، تمامی الگوریتم‌های FL در آموزش مدل‌های زبانی بزرگ نسبت به آموزش محلی بهبود قابل توجهی را ارائه می‌دهند. این به‌ویژه برای سازمان‌هایی که با داده‌های محدود مواجه هستند، یک انگیزه قوی برای مشارکت در یادگیری فدرال ایجاد می‌کند.

**2. مقایسه در سناریوهای داده ناهمگن**

* **FedAvg**: در شرایطی با داده‌های ناهمگن عملکرد متوسطی دارد و در برخی موارد نمی‌تواند به خوبی از تنوع داده‌ها بهره‌برداری کند.
* **FedProx**: در مقایسه با FedAvg، به‌طور قابل توجهی عملکرد بهتری در شرایط ناهمگن ارائه می‌دهد. این الگوریتم به‌خوبی می‌تواند فاصله بین مدل محلی و مدل جهانی را کنترل کند.
* **SCAFFOLD**: این الگوریتم نیز به‌ویژه در سناریوهای ناهمگن، به دلیل اصلاح گرادیان‌ها، دقت بهتری را ارائه می‌دهد.

**3. سرعت همگرایی**

* **FedAvgM** و **FedAdam**: این دو الگوریتم به دلیل بهبود در همگرایی، سرعت بالاتری در رسیدن به دقت مطلوب دارند. FedAvgM با استفاده از مؤلفه مومنتوم عملکرد پایدارتری دارد.
* **FedYogi**: با استفاده از روش‌های بهینه‌سازی پیشرفته، سرعت یادگیری را افزایش می‌دهد و می‌تواند به سرعت به نتایج بهتری دست یابد.

**4. عملکرد در حوزه‌های خاص (مانند مالی)**

در برخی آزمایشات، به‌ویژه در حوزه‌های خاصی مانند مالی، الگوریتم‌های FL توانسته‌اند عملکرد بهتری نسبت به مدل‌های مشهور مانند GPT-4 ارائه دهند. به‌طور خاص، FedProx و SCAFFOLD در این سناریوها عملکرد بهتری از خود نشان داده‌اند.

**5. نتایج کمی**

در آزمایشات تجربی، نویسندگان از بیش از 30 معیار ارزیابی برای سنجش عملکرد استفاده کرده‌اند. این معیارها شامل دقت، سرعت همگرایی، و ظرفیت مدل برای یادگیری از داده‌های توزیع‌شده است. به‌عنوان مثال، در آزمون MT-Bench، بهبود حداقل 12% در عملکرد مدل‌های آموزش‌دیده با FL نسبت به آموزش محلی مشاهده شده است.

**نتیجه‌گیری**

به‌طور کلی، الگوریتم‌های FL در OpenFedLLM به‌گونه‌ای طراحی شده‌اند که بتوانند در شرایط مختلف عملکرد مؤثری داشته باشند. انتخاب الگوریتم مناسب بسته به نوع داده‌ها و نیازهای خاص کاربرد می‌تواند تأثیر قابل توجهی بر نتایج نهایی داشته باشد.

### آزمایش

در بخش آزمایش، نویسندگان جزئیات مربوط به تنظیمات تجربی، مجموعه داده‌ها و روش‌های ارزیابی را ارائه می‌دهند. آنها به بررسی عملکرد الگوریتم‌های مختلف FL در زمینه‌های مختلف، از جمله مالی و پزشکی، می‌پردازند. نتایج نشان می‌دهد که آموزش از طریق FL به‌طور مداوم بهتر از آموزش محلی عمل می‌کند و در برخی موارد، مدل‌های آموزش‌دیده با FL قابلیت‌های بهتری نسبت به مدل‌های بزرگ شناخته‌شده دارند. نویسندگان همچنین به ارزیابی‌های مختلف و مجموعه داده‌های گوناگون اشاره می‌کنند که در آزمایشات خود استفاده کرده‌اند

بخش آزمایش‌ها در این مقاله به بررسی عملکرد الگوریتم‌های مختلف یادگیری فدرال (FL) در آموزش مدل‌های زبانی بزرگ (LLMs) می‌پردازد. این آزمایش‌ها شامل ارزیابی‌های تجربی متعدد با استفاده از مجموعه داده‌ها و معیارهای متنوع است. در اینجا جزئیات کلیدی این بخش آورده شده است:

#### 1. **پیکربندی تجربی**

* **مجموعه داده‌ها**: نویسندگان از هشت مجموعه داده مختلف برای آموزش و ارزیابی مدل‌ها استفاده کردند. این مجموعه داده‌ها شامل داده‌های عمومی و خصوصی هستند که در حوزه‌های مختلف مانند مالی، پزشکی و عمومی قرار دارند.
* **پروتکل یادگیری فدرال**: آزمایش‌ها بر اساس پروتکل‌های استاندارد FL انجام شده است که شامل مراحل زیر است:
  + بارگذاری مدل جهانی به مشتریان.
  + آموزش محلی مدل در هر مشتری با استفاده از داده‌های محلی.
  + ارسال مدل‌های محلی به سرور.
  + تجمیع مدل‌های محلی برای به‌روزرسانی مدل جهانی.

#### 2. **معیارهای ارزیابی**

برای ارزیابی عملکرد الگوریتم‌ها، نویسندگان از بیش از 30 معیار مختلف استفاده کرده‌اند. این معیارها شامل موارد زیر هستند:

* **دقت**: اندازه‌گیری توانایی مدل در پیش‌بینی نتایج صحیح.
* **سرعت همگرایی**: زمان مورد نیاز مدل برای رسیدن به سطح خاصی از دقت.
* **کارایی**: ارزیابی عملکرد مدل در شرایط مختلف و با داده‌های متفاوت.
* **توزیع‌پذیری**: بررسی توانایی مدل در یادگیری از داده‌های توزیع‌شده و ناهمگن.

#### 3. **نتایج تجربی**

* **بهبود در عملکرد**: نتایج نشان می‌دهد که تمامی الگوریتم‌های FL در مقایسه با آموزش محلی، بهبود قابل توجهی در عملکرد دارند. به‌خصوص، الگوریتم‌هایی مانند FedProx و SCAFFOLD در شرایط ناهمگن بهتر عمل کردند.
* **مقایسه با مدل‌های شناخته‌شده**: در برخی از آزمایش‌ها، مدل‌های آموزش‌دیده با FL توانسته‌اند عملکرد بهتری نسبت به مدل‌های مشهور مانند GPT-4 در حوزه‌های خاص ارائه دهند. به‌عنوان مثال، در آزمون‌های مالی، مدل‌های آموزش‌دیده با الگوریتم‌های FL به‌خصوص FedProx، عملکرد بهتری نسبت به GPT-4 نشان دادند.
* **زمان آموزش**: الگوریتم‌های FL توانستند آموزش را در مدت زمان کمتری نسبت به روش‌های سنتی انجام دهند. به‌طور متوسط، آموزش یک مشتری تنها به 1 تا 2 ساعت برای 100 دور ارتباط نیاز داشت.

#### 4. **تحلیل نتایج**

نویسندگان در این بخش نتایج را تحلیل کرده و به این نکته اشاره می‌کنند که هیچ یک از الگوریتم‌ها نمی‌تواند در تمام سناریوها بهترین عملکرد را ارائه دهد. به همین دلیل، انتخاب الگوریتم مناسب بسته به نوع داده‌ها و نیازهای خاص کاربرد بسیار مهم است. همچنین، نیاز به تحقیقات بیشتر برای توسعه الگوریتم‌های FL جدید و بهینه‌سازی تکنیک‌های موجود احساس می‌شود.

### نتیجه‌گیری

بخش آزمایش‌ها نتایج قابل توجهی را در زمینه بهبود عملکرد مدل‌های LLM از طریق یادگیری فدرال ارائه می‌دهد. این نتایج نشان‌دهنده پتانسیل بالای الگوریتم‌های FL در آموزش مدل‌های زبانی بزرگ و انگیزه قوی برای سازمان‌ها به منظور مشارکت در این نوع آموزش است.

### بحث و رهنمود آینده

در این بخش، نویسندگان به تحلیل نتایج تجربی و اهمیت آنها برای تحقیقات آینده می‌پردازند. آنها به چالش‌ها و فرصت‌های موجود در زمینه یادگیری فدرال و LLMها اشاره می‌کنند و به نیاز به توسعه الگوریتم‌های جدید FL برای بهبود عملکرد مدل‌های زبانی بزرگ تأکید می‌کنند. همچنین، آنها به این نکته اشاره می‌کنند که نیاز به همکاری بیشتر بین جوامع FL و LLM وجود دارد تا بتوانند از پیشرفت‌های یکدیگر بهره‌برداری کنند. در نهایت، نویسندگان به امید اینکه دیگران بتوانند بر اساس OpenFedLLM تحقیقات بیشتری را انجام دهند، تأکید می‌کنند.

**"Federated Learning Meets Natural Language Processing"**

### چکیده

چکیده مقاله به معرفی یادگیری فدرال (Federated Learning) به عنوان یک رویکرد نوین در یادگیری ماشین می‌پردازد که به ویژه برای حفظ حریم خصوصی داده‌های کاربران طراحی شده است. با توجه به اینکه بیشتر داده‌ها در دستگاه‌های کاربر نهایی ذخیره می‌شوند، یادگیری فدرال اجازه می‌دهد تا مدل‌های یادگیری ماشین بدون نیاز به انتقال داده‌های محلی به سرور مرکزی آموزش داده شوند. در این مقاله، نویسندگان به بررسی چالش‌های موجود در کاربرد یادگیری فدرال در پردازش زبان طبیعی (NLP) می‌پردازند و روش‌ها و ابزارهای ارزیابی موجود را نقد و بررسی می‌کنند. همچنین، خلأهای تحقیقاتی و مسیرهای آینده برای توسعه این حوزه را مورد بحث قرار می‌دهند.

### مقدمه

در مقدمه، نویسندگان به روندهای کنونی در یادگیری ماشین و به‌ویژه در پردازش زبان طبیعی اشاره می‌کنند. بیشتر تکنیک‌های NLP به داده‌های زیادی نیاز دارند که معمولاً از کاربران نهایی جمع‌آوری می‌شوند. به دلیل نگرانی‌های مربوط به حریم خصوصی، نیاز به رویکردهایی که داده‌های محلی را حفظ کنند و در عین حال مدل‌های کارآمدی را آموزش دهند، بیشتر احساس می‌شود. یادگیری فدرال به عنوان یک راه‌حل برای این مشکل معرفی می‌شود که به کاربران اجازه می‌دهد تا بدون به خطر انداختن اطلاعات شخصی خود، در فرآیند آموزش مدل‌ها شرکت کنند.

### یادگیری فدرال

این بخش به تشریح اصول بنیادی یادگیری فدرال می‌پردازد. یادگیری فدرال یک رویکرد توزیع‌شده است که در آن چندین دستگاه محلی به آموزش مدل‌ها پرداخته و تنها به اشتراک‌گذاری به‌روزرسانی‌های مدل با یک سرور مرکزی می‌پردازند. این رویکرد به طور خاص برای مقابله با چالش‌های داده‌های غیریکسان (non-IID) در محیط‌های واقعی طراحی شده است. مقاله به بررسی الگوریتم‌ها و چارچوب‌های مختلفی که برای یادگیری فدرال توسعه داده شده‌اند، می‌پردازد و انواع مختلف داده‌های توزیع‌شده را مورد بررسی قرار می‌دهد.

### یادگیری فدرال در NLP

در این بخش، به کاربردهای خاص یادگیری فدرال در پردازش زبان طبیعی پرداخته می‌شود. نویسندگان به پروژه‌های مختلفی اشاره می‌کنند که از یادگیری فدرال برای بهبود عملکرد مدل‌های NLP استفاده کرده‌اند، مانند پیش‌بینی کلمات در کیبوردهای موبایل و تشخیص کلمات کلیدی. همچنین، چالش‌های خاصی که در این زمینه وجود دارد، از جمله نیاز به حفظ حریم خصوصی و کیفیت پایین داده‌های توزیع‌شده، مورد بحث قرار می‌گیرد.

البته! در ادامه، به توضیح مفصل‌تری درباره‌ی **یادگیری فدرال در پردازش زبان طبیعی (NLP)** می‌پردازم:

### یادگیری فدرال در NLP

یادگیری فدرال به عنوان یک رویکرد نوآورانه در یادگیری ماشین، به ویژه در حوزه پردازش زبان طبیعی (NLP) کاربردهای قابل توجهی دارد. این رویکرد به دلیل مزایای خاصی که در حفظ حریم خصوصی و بهبود دقت مدل‌ها ارائه می‌دهد، در حال جذب توجه زیادی از سوی محققان و توسعه‌دهندگان است.

#### 1. **مزایای یادگیری فدرال در NLP**

* **حفظ حریم خصوصی:** یکی از بزرگ‌ترین مزایای یادگیری فدرال در NLP این است که داده‌های حساس کاربران (مانند متن‌های پیام، جستجوها و سایر تعاملات) به طور محلی در دستگاه‌های کاربر باقی می‌مانند و هیچ‌گونه اطلاعاتی به سرور مرکزی منتقل نمی‌شود. این امر به ویژه در زمینه‌هایی مانند بهداشت و درمان یا خدمات مالی که اطلاعات شخصی بسیار حساس است، اهمیت دارد.
* **استفاده از داده‌های توزیع‌شده:** داده‌های NLP معمولاً از منابع مختلفی جمع‌آوری می‌شوند و ممکن است در دستگاه‌های مختلف توزیع شده باشند. یادگیری فدرال این امکان را فراهم می‌کند که از این داده‌های توزیع‌شده برای آموزش مدل‌ها استفاده شود بدون اینکه نیازی به جمع‌آوری آن‌ها در یک مکان مرکزی باشد.

#### 2. **کاربردهای یادگیری فدرال در NLP**

* **پیش‌بینی کلمات در کیبوردهای موبایل:** یکی از کاربردهای بارز یادگیری فدرال در NLP، بهبود عملکرد کیبوردهای هوشمند مانند Gboard است. در این سیستم‌ها، مدل‌های NLP به طور محلی بر روی داده‌های تایپ کاربر آموزش می‌بینند و پیش‌بینی‌های بهتری از کلمات و عبارات ارائه می‌دهند، بدون اینکه داده‌های حساس به سرور ارسال شوند.
* **شخصی‌سازی خدمات:** یادگیری فدرال می‌تواند به شخصی‌سازی نتایج جستجو، پیشنهادات محتوا و سیستم‌های پرسش و پاسخ کمک کند. با استفاده از داده‌های محلی کاربران، می‌توان مدل‌هایی را آموزش داد که به طور خاص بر اساس رفتار و ترجیحات فردی کاربر عمل کنند.
* **تشخیص گفتار:** در سیستم‌های تشخیص گفتار، یادگیری فدرال می‌تواند به بهبود دقت و سرعت مدل‌ها کمک کند. داده‌های صوتی معمولاً در دستگاه‌های کاربر ذخیره می‌شوند و یادگیری فدرال می‌تواند به آموزش مدل‌هایی بپردازد که به طور مؤثری از این داده‌ها استفاده کنند.

#### 3. **چالش‌های یادگیری فدرال در NLP**

* **داده‌های غیریکسان:** یکی از چالش‌های اصلی در یادگیری فدرال، عدم تعادل و غیریکسانی داده‌ها در دستگاه‌های مختلف است. به عنوان مثال، ممکن است یک کاربر تعداد زیادی متن مربوط به یک موضوع خاص داشته باشد در حالی که کاربر دیگری فقط اطلاعات محدود و غیرمرتبط داشته باشد. این عدم تعادل می‌تواند بر کیفیت یادگیری تأثیر بگذارد.
* **حفظ حریم خصوصی و امنیت:** اگرچه یادگیری فدرال به طور طبیعی حریم خصوصی را حفظ می‌کند، اما هنوز خطرات امنیتی وجود دارد. به اشتراک‌گذاری به‌روزرسانی‌های مدل ممکن است منجر به افشای اطلاعات حساس شود، بنابراین نیاز به تکنیک‌های پیشرفته‌ای برای حفظ امنیت اطلاعات است.
* **محدودیت‌های محاسباتی:** دستگاه‌های کاربر معمولاً دارای محدودیت‌های پردازشی و حافظه‌ای هستند. این موضوع می‌تواند موجب کاهش کارایی مدل‌های یادگیری فدرال شود، به ویژه در شرایطی که نیاز به پردازش داده‌های پیچیده وجود دارد.

#### 4. **آینده یادگیری فدرال در NLP**

آینده یادگیری فدرال در پردازش زبان طبیعی بسیار امیدوارکننده است. با پیشرفت در تکنیک‌های یادگیری عمیق و الگوریتم‌های یادگیری فدرال، می‌توان انتظار داشت که این رویکرد به تدریج در کاربردهای مختلف NLP، از جمله ترجمه ماشینی، تحلیل احساسات و سیستم‌های گفتاری پیشرفته، به کار گرفته شود. همچنین، توسعه ابزارها و روش‌های جدید برای ارزیابی و بهینه‌سازی مدل‌ها در محیط‌های توزیع‌شده می‌تواند به بهبود عملکرد این سیستم‌ها کمک کند.

### نتیجه‌گیری

یادگیری فدرال در پردازش زبان طبیعی نه تنها به حفظ حریم خصوصی کاربران کمک می‌کند، بلکه امکان بهره‌برداری از داده‌های توزیع‌شده را نیز فراهم می‌سازد. با وجود چالش‌های موجود، این رویکرد پتانسیل بالایی برای بهبود عملکرد مدل‌های NLP دارد و می‌تواند به توسعه خدمات شخصی‌سازی‌شده و هوشمند کمک کند.

### معیارها

این بخش به معیارهای ارزیابی مدل‌های یادگیری فدرال می‌پردازد. نویسندگان به بررسی روش‌های مختلفی که برای ارزیابی عملکرد مدل‌ها در محیط‌های توزیع‌شده به کار می‌روند، می‌پردازند. معیارهایی همچون دقت، صحت و کارایی در این راستا مورد بحث قرار می‌گیرند. همچنین، به چالش‌هایی که در ارزیابی مدل‌ها در شرایط غیرمتمرکز وجود دارد، اشاره می‌شود.

البته! در ادامه، به توضیح مفصل‌تری درباره‌ی **معیارها** (Metrics) در یادگیری فدرال و پردازش زبان طبیعی (NLP) می‌پردازم:

### معیارها در یادگیری فدرال و پردازش زبان طبیعی

معیارها نقش اساسی در ارزیابی عملکرد مدل‌های یادگیری ماشین دارند، به ویژه در زمینه یادگیری فدرال که داده‌ها به صورت توزیع‌شده و محلی ذخیره می‌شوند. این معیارها به ما کمک می‌کنند تا کیفیت مدل‌ها را بسنجیم و بهینه‌سازی‌های لازم را انجام دهیم.

#### 1. **معیارهای عمومی ارزیابی مدل**

* **دقت (Accuracy):** دقت یکی از ساده‌ترین و پرکاربردترین معیارها برای ارزیابی مدل‌های یادگیری ماشین است. این معیار نشان‌دهنده درصد پیش‌بینی‌های صحیح مدل نسبت به کل پیش‌بینی‌ها است. با این حال، دقت ممکن است در موقعیت‌هایی که داده‌ها غیریکسان هستند، اطلاعات کافی ندهد.
* **حساسیت و اختصاصیت (Sensitivity and Specificity):** حساسیت (Recall) نشان‌دهنده توانایی مدل در شناسایی نمونه‌های مثبت است، در حالی که اختصاصیت (Precision) نشان‌دهنده توانایی مدل در شناسایی نمونه‌های منفی است. این دو معیار به ویژه در مسائل دسته‌بندی با کلاس‌های نامتعادل اهمیت زیادی دارند.
* **F1 Score:** این معیار ترکیبی از حساسیت و اختصاصیت است و در شرایطی که تعادل بین این دو مهم است، به کار می‌رود. F1 Score به طور خاص در مسائلی که هزینه‌های متفاوتی برای خطاهای مثبت و منفی وجود دارد، مفید است.

#### 2. **معیارهای خاص یادگیری فدرال**

* **همگرایی مدل (Model Convergence):** در یادگیری فدرال، همگرایی مدل به معنای رسیدن به یک نقطه بهینه است که در آن به‌روزرسانی‌های مدل دیگر تغییرات معناداری ایجاد نمی‌کنند. ارزیابی همگرایی می‌تواند شامل بررسی تعداد دورهای آموزش (Epochs) و میزان تغییرات در عملکرد مدل باشد.
* **هزینه ارتباطی (Communication Cost):** یکی از چالش‌های اصلی در یادگیری فدرال، هزینه‌های مرتبط با ارتباط بین دستگاه‌های محلی و سرور مرکزی است. این هزینه‌ها می‌توانند شامل زمان و منابع مصرفی برای ارسال به‌روزرسانی‌های مدل باشند. معیارهایی برای اندازه‌گیری این هزینه‌ها می‌توانند شامل حجم داده‌های ارسال‌شده و زمان مورد نیاز برای انتقال اطلاعات باشند.
* **میزان مشارکت کاربران (User Participation Rate):** در یادگیری فدرال، میزان مشارکت کاربران در هر دور آموزش می‌تواند بر عملکرد کلی مدل تأثیر بگذارد. این معیار نشان‌دهنده درصد کاربرانی است که در فرآیند آموزش شرکت می‌کنند و می‌تواند بر اساس نوع داده‌ها و نیازهای مدل تغییر کند.

#### 3. **معیارهای ارزیابی خاص NLP**

* **نمره BLEU (Bilingual Evaluation Understudy):** این معیار به‌ویژه در ترجمه ماشینی برای ارزیابی دقت ترجمه‌ها به کار می‌رود. نمره BLEU با مقایسه خروجی مدل با ترجمه‌های مرجع محاسبه می‌شود و به ما کمک می‌کند تا کیفیت ترجمه‌ها را اندازه‌گیری کنیم.
* **نمره ROUGE (Recall-Oriented Understudy for Gisting Evaluation):** این معیار برای ارزیابی کیفیت خلاصه‌سازی متون استفاده می‌شود. ROUGE با مقایسه کلمات و عبارات کلیدی در متن‌های تولیدشده و مرجع، کیفیت خلاصه‌ها را اندازه‌گیری می‌کند.
* **نمره PERPLEXITY:** در مدل‌های زبانی، نمره perplexity به عنوان معیاری برای ارزیابی کیفیت پیش‌بینی‌های مدل استفاده می‌شود. نمره پایین‌تر به معنای کیفیت بهتر مدل در پیش‌بینی توالی‌های کلمات است.

#### 4. **چالش‌های ارزیابی در یادگیری فدرال**

* **داده‌های غیریکسان:** در یادگیری فدرال، داده‌ها معمولاً توزیع‌شده و غیریکسان هستند. این موضوع می‌تواند منجر به بروز مشکلاتی در ارزیابی دقیق عملکرد مدل شود، به‌ویژه زمانی که داده‌ها به صورت نامتعادل در دستگاه‌های مختلف توزیع شده‌اند.
* **تأثیرات حریم خصوصی:** در یادگیری فدرال، به دلیل محدودیت‌ها در به اشتراک‌گذاری داده‌ها، ممکن است در برخی موارد نتوان معیارهای ارزیابی را به طور کامل محاسبه کرد. به همین دلیل، نیاز به توسعه معیارهای جدید و سازگار با شرایط یادگیری فدرال احساس می‌شود.

### نتیجه‌گیری

معیارها ابزارهای حیاتی برای ارزیابی عملکرد مدل‌های یادگیری فدرال در پردازش زبان طبیعی هستند. با توجه به چالش‌های خاص این حوزه، توسعه و بهینه‌سازی معیارهای ارزیابی می‌تواند به بهبود عملکرد مدل‌ها و افزایش کارایی سیستم‌های NLP کمک کند. این امر نه تنها در زمینه یادگیری فدرال بلکه در کل حوزه یادگیری ماشین نیز دارای اهمیت است.

### چالش و مسیر آینده

در این بخش، نویسندگان به چالش‌های باقی‌مانده در زمینه یادگیری فدرال و پردازش زبان طبیعی اشاره می‌کنند. مشکلاتی از جمله کارایی الگوریتم‌ها در شرایط غیرمتمرکز، نیاز به بهبود تکنیک‌های حریم خصوصی و کمبود داده‌های باکیفیت در بسیاری از کاربردها مطرح می‌شود. نویسندگان همچنین به مسیرهای تحقیقاتی آینده که می‌توانند به بهبود یادگیری فدرال و کاربردهای آن در NLP کمک کنند، می‌پردازند.

یادگیری فدرال (Federated Learning) با چالش‌های خاصی در زمینه داده‌های کم‌حجم مواجه است. در زیر به برخی از این چالش‌ها پرداخته می‌شود:

### 1. ****عدم تعادل داده‌ها****

در یادگیری فدرال، داده‌ها معمولاً در دستگاه‌های مختلف توزیع شده‌اند و ممکن است هر دستگاه فقط مقدار کمی از داده‌ها را داشته باشد. این عدم تعادل می‌تواند منجر به skewness (عدم توازن) در توزیع داده‌ها شود که به نوبه خود می‌تواند بر عملکرد مدل تأثیر منفی بگذارد. به عنوان مثال، اگر یک دستگاه تنها داده‌های مربوط به یک کلاس خاص را داشته باشد، مدل ممکن است به خوبی آموزش نبیند و دقت آن در پیش‌بینی‌های مربوط به کلاس‌های دیگر کاهش یابد.

### 2. ****چالش‌های یادگیری در شرایط غیریکسان (Non-IID)****

داده‌های کم‌حجم غالباً غیریکسان هستند، به این معنی که توزیع داده‌ها در دستگاه‌های مختلف می‌تواند به شدت متفاوت باشد. این چالش در یادگیری فدرال می‌تواند به بروز مشکلاتی در همگرایی مدل منجر شود. به عنوان مثال، اگر برخی از دستگاه‌ها داده‌های متنوع‌تری داشته باشند و برخی دیگر داده‌های محدودی، مدل نهایی ممکن است نتواند به خوبی برای همه انواع داده‌ها تعمیم یابد.

### 3. ****مدل‌های پیچیده و محدودیت‌های محاسباتی****

دستگاه‌های کاربر نهایی معمولاً دارای محدودیت‌های سخت‌افزاری هستند. این محدودیت‌ها می‌تواند شامل توان پردازشی پایین، حافظه محدود و نیاز به پاسخگویی سریع باشد. در شرایطی که داده‌های کم‌حجم وجود دارد، این محدودیت‌ها می‌توانند عملکرد مدل را تحت تأثیر قرار دهند و منجر به یادگیری ناکافی یا بیش‌برازش (overfitting) شوند.

### 4. ****مسائل حریم خصوصی و امنیت****

حفظ حریم خصوصی در یادگیری فدرال اهمیت ویژه‌ای دارد، به ویژه زمانی که داده‌های کم‌حجم و حساس در دسترس است. در این شرایط، اگرچه داده‌ها به طور محلی باقی می‌مانند، اما به اشتراک‌گذاری مدل‌های یادگیری و به‌روزرسانی‌های آن‌ها می‌تواند اطلاعات حساس را فاش کند. این مسأله به ویژه در صنایعی مانند بهداشت و درمان و مالی با توجه به حساسیت اطلاعات، اهمیت بیشتری پیدا می‌کند.

### 5. ****قابلیت تعمیم****

مدل‌های یادگیری فدرال نیاز به تعمیم‌پذیری خوبی دارند تا بتوانند در شرایط واقعی و متنوع کار کنند. با داده‌های کم‌حجم، ممکن است مدل‌ها به خوبی تعمیم نیابند و در شرایط واقعی دچار مشکل شوند. این مسأله می‌تواند منجر به کاهش دقت و کارایی مدل‌ها در پیش‌بینی‌های جدید شود.

### 6. ****نیاز به الگوریتم‌های بهینه‌سازی پیشرفته****

در شرایط داده‌های کم‌حجم و توزیع‌شده، نیاز به استفاده از الگوریتم‌های بهینه‌سازی پیشرفته و مناسب وجود دارد. این الگوریتم‌ها باید به گونه‌ای طراحی شوند که بتوانند با داده‌های محدود و غیریکسان کار کنند و همزمان حریم خصوصی کاربران را حفظ کنند.

چالش‌های یادگیری فدرال در زمینه داده‌های کم‌حجم نیاز به توجه ویژه و تحقیق بیشتر دارد. توسعه الگوریتم‌ها و تکنیک‌های جدید که بتوانند این چالش‌ها را برطرف کنند، کلید موفقیت در به کارگیری یادگیری فدرال در محیط‌های واقعی است.

### نتیجه‌گیری

در نتیجه‌گیری، نویسندگان بر اهمیت یادگیری فدرال در عصر داده‌های توزیع‌شده و نیاز به حفظ حریم خصوصی تأکید می‌کنند. آن‌ها به مزایای بالقوه این رویکرد در بهبود عملکرد مدل‌های NLP و کاهش خطرات امنیتی اشاره می‌کنند. همچنین، بر لزوم تحقیقات بیشتر و توسعه ابزارهای جدید برای بهبود یادگیری فدرال در زمینه پردازش زبان طبیعی تأکید می‌شود.