سلام عنوان مقاله من هست اٍ کامپری هٍنسیو ری وی یو آن اِين سِمبٍل دیپ لرنینگ آپرچٍنِتیز اًند چًلِنجِز

که ترجمش میشه : مروری جامع بر یادگیری عمیق گروهی: فرصت ها و چالش ها

نویسندگان این مقاله عمار محمد و رانیا کورا از گروه کامپیوتر دانشگاه کایرو از کشور مصر هستند

و این مقاله در تاریخ 19 جًنییوری 2023 اکسپت شده و در تاریخ 1 فٍبریه 2023 به صورت آنلاین در دسترس عموم قرار گرفته

لینک دی او آی مقاله رم این قسمت برای شما قرار دادم

اسلاید 2

امروزه بحث یادگیری ماشین توی دنیایی که پر از منابع متفاوت و گوناگون شده به یکی از مهمترین و قالبترین شاخه های روش های هوش مصنوعی تبدیل شده که در بسیاری از حوزه ها به کار برده میشه، خوب انواع الگوریتم های یادگیری و روش های مختلفی وجود داره که نقاط ضعف و معایب هر روش بر اساس چندین عامل از جمله عملکرد و مقیاس پذیری اندازه گیری میشن بر اساس تحقیقات گسترده که در زمینه یادگیری ماشین صورت گرفته، دو روش بر الگوریتم‌های یادگیری ماشین تسلط دارن1- یادگیری عمیق 2- یادگیری جمعی یا ترکیبی.

تکنیک‌های یادگیری عمیق , می‌تونن روی مباحثی مثل مقیاس‌پذیری و حل مسائل پیچیده بپردازند و استخراج ویژگی‌های خودکار از داده‌های بدون ساختار رو ارائه بدن. علاوه بر این، روش‌های یادگیری عمیق شامل چندین نوع معماری شبکه برای وظایف مختلفی هستند، مثل شبکه‌های عصبی پیش‌خور، شبکه‌های عصبی کانولوشنی و شبکه‌های عصبی بازگشتی. اما مشکل اینجاست که فرآیند آموزش مدل‌های یادگیری عمیق نیاز به تلاش بسیار زیادی و طاقت فرسایی داره و تنظیم بهینه پارامترهاش نیازمند تخصص و آزمایش‌های گسترده ای که به نوبه خودش وظیفه‌ای زمان‌بر و خسته‌کنندس.

اسلاید 3

از طرف دیگه، یادگیری جمعی به یک روش یادگیری اشاره داره که چندین مدل پایه رو ترکیب می‌کنه، تکرار میکنم چندین مدل پایه رو ترکیب می کنه تا یک مدل بزرگ‌تر و قدرتمندتر ایجاد کنه. علاوه بر این، یادگیری جمعی می‌تونه به دلیل تنوع مدل‌های پایه، خطر.... های واریانس یا (overfitting) رو کاهش بده. یادگیری جمعی به طور موفقیت‌آمیزی در حوزه‌ها و زمینه‌های مختلف اعمال شده و عملکرد بهتری نسبت به مدل‌های تکی داره. تکنیک‌های مختلفی برای یادگیری جمعی وجود داره که در نحوه آموزش و ترکیب مدل‌های پایه با همدیگه تفاوت دارن. رایج‌ترین تکنیک‌های یادگیری جمعی شامل میانگین‌ گیری (averaging)، بگینگ (bagging)، استکینگ (stacking) و بوس تینگ (boosting) هستند که توی اسلاید های بعدی تک تک این مدل ها رو که بحث اصلی مقاله ما هستند دقیقا با هم بررسی می کنیم.

بنابراین، این مقاله هدف داره تا به‌طور جامع استراتژی‌های مختلف برای به‌کارگیری یادگیری عمیق جمعی یا (deep ensemble learning) رو بررسی کنه. همچنین، جنبه‌های مختلفی رو که بر موفقیت روش‌های جمعی تأثیر میزارن، از جمله نوع مدل‌های پایه استفاده شده، تکنیک‌های نمونه‌گیری داده‌ها در آموزش، تنوع به‌کارگیری طبقه‌بندهای پایه مختلف و روش‌های ترکیب مدل‌های عمیق پایه، ارائه کنه. همچنین مزایای هر استراتژی رو مورد بحث قرار می‌دهد.

-اسلاید 4

به دلیل قدرت و کارایی سیستم یادگیری جمعی در بهبود عملکرد پیش‌بینی مدل، یادگیری جمعی به یک روند تحقیقاتی مهم در سال‌های اخیر تبدیل شده ، که منجر به افزایش مطالعات که از یادگیری جمعی در حوزه‌های مختلف کاربردی استفاده می‌کنن شده. به منظور نشان دادن این که در هر سال به چه اندازه تعداد مقالات منتشر شده درباره یادگیری جمعی افزایش یافته و در حوزه‌های مختلفی که یادگیری جمعی در آنها اعمال شده از سال 2014 تا 2021، یک جستجو در پایگاه داده اسکوپس با استفاده از اصطلاحات "یادگیری جمعی" و "یادگیری عمیق جمعی" توی این مقاله انجام شده. این اصطلاحات در عناوین، چکیده‌ها و کلمات کلیدی مقالات جستجو شدن.

طبق آمار به دست اومده از این جستوجو به نظر می‌رسه که یادگیری جمعی و یادگیری عمیق جمعی در حال رشد قابل توجهی هستند، به ویژه در حوزه علوم کامپیوتر. آمارهای ارائه شده نشون میدن که این روش‌ها به دلیل توانایی‌ که در بهبود عملکرد پیش‌بینی دارن، توجه محققان زیادی رو به خودشون جلب کردن. به دلیل اینکه بیشترین تعداد مقالات مرجع به یادگیری جمعی و یادگیری عمیق جمعی در سال 2021 منتشر شده ، نشون میده که پژوهشگران به فعالیت و استفاده از این روش‌ها در حوزه‌های مختلف علاقه‌مند هستند. همچنین نرخ استفاده بالا در علوم کامپیوتر نشون دهنده اینکه این تکنیک‌ها در این حوزه اهمیت و کاربرد عملی قابل توجهی دارند.

اسلاید 5

همونطور که ملاحظه می فرمایید شکل ۱ تعداد مقالات منتشر شده برای اصطلاح جستجوی "یادگیری جمعی" در هر سال در طول دوره مورد بحث ماست که از سال 2014 تا 2021 هست این شکل نشون میده تعداد مقالاتی که با استفاده از این اصطلاح پیدا شده‌اند، تقریباً ۲۵,۲۶۲ مقاله است، که نشان از روند افزایشی یادگیری جمعی در طی چند ساله داره.

اسلاید 6

شکل ۲ تعداد مقالاتی که به اصطلاح جستجوی "یادگیری جمعی" در تمامی حوزه‌ها پرداخته‌ شده رو به ما نشون میده . از این شکل می‌تونیم مشاهده کنیم که حوزه علوم کامپیوتر دارای بیشترین تعداد مقالهاست، که تقریباً ۱۶,۷۸۲ مقاله تخمین زده شده.

اسلاید 7

در مقابل بحث یادگیری جمعی شکل ۳ تعداد مقالات منتشر شده برای اصطلاح جستجوی "یادگیری عمیق جمعی" در هر سال در طول دوره مورد بحث ماست. تعداد مقالاتی که با استفاده از این اصطلاح پیدا شده‌، تقریباً ۶,۱۷۳ مقاله است، که نشان از علاقه روزافزون پژوهشگران به این موضوع هستش.

اسلاید 8

شکل ۴ تعداد مقالاتی که به اصطلاح جستجوی "یادگیری عمیق جمعی" در تمام حوزه‌ها پرداخته‌ شده رو نمایش میده. از این شکل می‌تونیم مشاهده کنیم که حوزه علوم کامپیوتر دارای بیشترین تعداد مقالهاست.

اسلاید 9

خوب توی این اسلاید میریم سراغ بحث اصلیمون که یادگیری جمعی یا(اٍنسمبٍل لرنینگ ensemble learning) هست گفتیم یکی دیگر از روش‌های یادگیری ماشین یادگیری جمعی هست که با در نظر گرفتن چندین مدل، دقت و عملکرد نهایی مدل را بهبود می‌بخشه. هدف از این روش یادگیری ماشین، کاهش خطاها و های بایاس که ممکنه در هر یک از مدل‌های تکی یادگیری ماشین وجود داشته باشه. در این حوزه، چندین مدل به صورت ترکیبی استفاده می‌شن تا توانایی مدل در پیش‌بینی داده‌های خروجی رو بهبود بدن.

اسلاید بعد

قبل از اینکه وارد بحث مدل های مورد برسسی در این مقاله بشیم یک مثال ساده برای نمایش عنواع خطاها توی بحث مدل های یادگیری ماشین برای درک بهتر موضوع رو مورد برسسی قرار میدم تا در طول ارائه کامل متوجه مباحث بشین می خوایم با هم عنواع خطاها رو مرور کنیم.

تابع اول نشون دهنده یک خط صاف در فضای داده مون هست که قراره مرز بین دوتا داده رو از هم تشخیص بده مثلا اگر مدل ما اس وی ام باشه و تابع hx ما این باشه استفاده از چنین hx نشون میده که این خط خیلی مقرور هست و همون داده های ترینیگم که در طول آموزش میبینه نمیتونه مدل کنه و انعطاف پذیری خیلی پایینی داره در چنین حالتی می گیم مشکل های بایاس داریم.

 حالا اگر ما تو مثال این تابع که مربوط به مدل پولی نومیال ما میشه مثلا یک درجه تنظیم براش بزاریم مثلا D رو درجه پولی نومیال در نظر بگیریم به اینصورت که بتونیم از 0 تا 10 مقدار تنظیم براش ست کنیم که اگر بالا باشه انعطاف پذیریش نسبت به مدل کردن داده های تست بیشتر یعنی راحت میتونه مدل کنه و اگر پایین باشه دشوارتر میشه مثل همون خط صاف یا اس وی اممون خوب حالا بیایم مثلا دی رو روی 10 تنظیم کنیم

چنین اچ ایکسی میتونه به خوبی داده های ترینیگ ست رو تخمین بزنه ولی توی این حالت هم روی داده های جدید نمیتونه مانور خوبی بده و مشکل تمیم پذیری داره یا همون جنریلیشن.

اسلاید 10

توی این اسلاید به یکی از مهمترین عوامل که می تونه بر عمکلرد یک سیستم همکاری جمعی تاثیر گذار باشه میپردازیم که اونم روش نمونه برداری داده هاست. . ما دو روش یا بهتره بگم دو استراتژی برای نمونه برداری داده ها داریم

1. استاتژی مجموعه داده های مستقل یا (ایٍندیپٍندٍنت دیتاست استراتزی independent datasets strategy) : در این استراتژی، زیرمجموعه‌ها به هم وابسته نیستند. این به این معناست که عملکرد یک زیرمجموعه بر عملکرد سایر زیرمجموعه‌ها تأثیر نمیزاره



اسلاید 11

۲- دومین استراتژی مجموعه داده‌های وابسته(دیپٍندٍنتdependent datasets strategy): برعکس استراتژی قبلی در این استراتژی، زیرمجموعه‌ها به یکدیگر وابسته هستند. این به این معناست که عملکرد یک زیرمجموعه تحت تأثیر نتایج زیرمجموعه‌های قبلی قرار می‌گیره.



اسلاید 12

دومین عامل موثر در یک سیستم گروهی تنوع مدل‌های پایه، هست ببینید اینجا مدل رو داریم بحث می کنیم نه دیتا ستمون. در مرکز هر سیستم مبتنی بر گروه، دو تکنیک برای آموزش مدل های گروه به طور جداگانه وجود دارد: تکنیک گروه توالی و تکنیک گروه موازی.

تکنیک گروه متوالی (ترتیبی یا سٍکوئنشٍل گروپ تکنیک Sequential Group Technique): توی این تکنیک مدل هامون به ترتیب آموزش داده می‌شن به همین دلیل وابستگی به داده وجود داره. بنابراین، اشتباهات مدل اول به ترتیب توسط مدل دوم اصلاح می‌شن، همونطور که در شکل ۷ دیدیم. مزیت اصلی روش‌های توالی اینه که از وابستگی بین مدل های پایه بهره مند میشن.

تکنیک گروه موازی(غیر ترتیبی یا پارالٍل Parallel Group Technique): در این تکنیک، مدل های یادگیری پایًمون به صورت همزمان ایجاد می‌شن و شروع به کار میکنن چون هیچ وابستگی به داده وجود نداره ، مثل شکل ۸. مزیت اصلی این تکنیک اینکه از استقلال بین مدل های پایه بهره می‌برن. بنابراین، اشتباهات تولید شده توسط یک مدل با اشتباهات موجود در یک مدل مستقل دیگه متفاوته، که به سیستم گروهی اجازه میدن که اشتباهات رو به صورت میانگین محاسبه کنن که راجب بحث میانگین گیری توی اسلاید های بعدیمون کامل با مثال صحبت می کنیم.

اسلاید 13

توی این جدول دسته‌بندی روش‌های گروهی که در مقاله مورد بررسی قرار گرفته‌اند، آورده شده است که ما می خوایم راجب تکنیک های استکینگ ، بگینگ و بوستینگ کامل بحث کنیم تا آشنایی کامل نسبت به این مدل ها پیدا کنیم



اسلاید 14

یادگیری ماشین با استفاده از روش استکینگ قبل از اینکه بخوایم وارد بحث مدل استکینگ بشیم بریم یه (کلیپ ) برای درک بیشتر موضوع با هم ببینیم

همونطور که در کلیپ دیدید کار گروهی خیلی بهتر از کار تکی هستش این موضوع توی مدل های دسته بندیمون هم صدق می کنه با توجه به بحث کلاس بندی که آقای دکتر غفاری هم سر کلاس گفتن که شامل یادگیری با ناظر یا نظراتی و نیمه نظارتی و یا مستقل بود ببینید مثلا اگر ما بخواییم دو مدل داده رو از دیتاستمون جدا کنیم باید از مدل های یادگیری ماشین استفاده کنیم که این مدل ها باید داده ها رو از هم جدا کنند این مدل ها به تنهایی نمی تونن کامل همه داده هامون رو برچسب دقیق بزنن ما قراره توی این مقاله سراق تکنیک هایی بریم که می توونن با استفاده از چندین مدل مثل مدل های اس وی ام و کا ان ان و یا درختی و ترکیب اونها با هم بتونن لیبل یا کلاس بندی رو توی بحث آموزش مدلمون بهتر انجام بدن.

خوب اول از همه بریم سراغ یک مثال ساده از تکنیک Voting و بعد بریم سراغ استکینک

هدف از این مثال آشنایی شما با موضوع مورد بحث هست در نظر بگیرید که مثلا شما توی دیتا بیس آموزشیتون یک دسته دیتا دارین که توی این دسته دیتا عکسهایی از سگها و گربها وجود داره که قرار یک تکنیکی بیاد این دیتا رو به ماشین آموزش بده تا بتونه این دو رو لیبل بزنه و کلاس بندی کنه یعنی چی یعنی بیاد توی دیتا بیس آموزشی که توش دو دسته دیتا از سگ ها و گربه ها وجود دارن رو تشخیص بده و عکس هایی که مرتبط به کلاس گربها هستند بر اساس ویژگیشون ببره تو کلاس گربه ها و دقیقا همینکار رو برای عکس هایی که مرتبط به کلاس سگها هستن انجام بده. توی یادگیری جمعی میاد این کار رو با چندین مدل انجام میده این کار باعث میشه که دقیقتر این مدل هامون دسته بندی ها رو تشخیص بدن و مشکلات رایج مثل های بایاس و های واریانس یا اٌور فیتینگ به حداقل برسونن . شما فرض کنین که قراره ما این کار رو مثلا با svm خطی بدون روش های یادگیری جمعی انجام بدیم یعنی بیاییم یک خط مستقیم رسم کنیم تا از وسطٍ دیتای آموزشیمون بگذره ....

و بعد بگیم هر دیتایی که سمت بالای خط بود لیبل سگ ها رو بخوره و هر دیتایی که پایین خط بود لیبل گربه ها رو بخوره خوب svm نمی تونه دقیق این کار رو به تنهایی انجام بده چون میاد فقط یک خط بین داده ها میکشه و ممکنه به اشتباه چنتا دیتا از سگ ها توی دسته بندی گربه ها بیفته که این عمل باعث بروز مشکل های بایاس میشه برای این مشکل اومدن از یادگیری جمعی استفاده کردن به چه شکل مثلا توی

 تکنیک VOITING ما یک دیتا بیس آموزشی داریم که میاد این دیتا رو به چندین مدل،همزمان تحویل میده مثلا یکی از مدل هامون svm یکی کا ان ان و یکی درختی بعد این مدل ها میان بر اساس ویژگی هایی که دارن این دیتا رو کلاس بندی میکنن و لیبل میزنن و در آخر همه جواب های مدل ها میاد توی قسمت تجمیع جمع میشه

 و در این قسمت بر اساس روش های تجمیع تکرار میکنم روش های تجمیع که توی اسلاید های قبل هم گفتیم مفصل راجبشون بحث میکنیم به هر مدل امتیاز میده......

 مثلا سه مدل داریم که دوتاشون گفتن این دیتا مرتبط به سگ هست یک مدل گفته این دیتا مرتبط به کلاس گربه ها هست (صبر تا انجام بدم)

 توی روش voiting این کار بر اساس بیشترین امتیاز انجام میشه مثلا شما در نظر بگیرید که ما دیتا رو تحویل 3 مدل دادیم یک مدل گفته این عکس مرتبط به سگ هاست یک مدل مثلا اس وی ام به اشتباه گفته این عکس مرتبط به گربه هاست و مدل سومم گفته این عکس مرتبط به سگ هاست در آخر توی تجمیع max voting میاد میگه خیله خوب هرچقدر امتیاز مدل ها به سگ ها یا گربه ها بیشتر باشه لیبل همون امتیاز بیشتر رو میخوره اینجا چون دوتا مدل تشخیص دادن این داده مرتبط به کلاس سگ هاست پس لیبل سگ ها روش می خوره . پس چی شد ما کارمون سه قسمت داره یکی دیتا ست آموزشیمون قسمت دوم مدل های آموزشیمون که می تونیم توی روش ویتینگ چندین مدل رو داشته باشیم مثلا یکی svm یکی کا ان ان و یکی درختی در قسمت سوم فرایند تجمیع صورت میگیره که چند روش تجمیع داریم یکیش max voting هست که الان توضیحش دادیم که مثلا برای مسائل لیبل بندی خیلی به درد می خوره گرچه که می تونیم توی مسائل رگراسیون هم ازش استفاده کنیم مثلا بگیم بین صفر و یک به هر مدلمون بر اساس تواناییش توی حل مسئله یک امتیاز بده و بعد بیایم امتیاز ها رو با هم جمع ببندیم و هر چقدر امتیاز سگها بیشتر بود توی چندین مدل بگیم بیاد لیبل سگها رو بزنه

. یکی دیگه از روش های تجمیع میانگین وزنی در مسائل رگرسوین هست و یکی هم averaging voting هست که میاد میانگین گیری میکنه اما بدون وزن. پس این شد تکنیک های یادگیری جمعی یا Ensmbel Learning با استفاده از روش voting

....

توی تکنیک استکینیک هر مدل بر اساس توانایی خودش سهم متفاوتی داره تقریبا شبیه تکنیک ویتنیگ هست با این تفاوت که هر کدوم از مدل ها میاد یک وزن میگیره یا در قسمت تجمیع رای گیری رو بر اساس وزن ها انجام میده که ایراد مدل ویتنیگ رو برطرف میکنه توی تکنیک ویتینگ گفتیم هر مدل سهم یکسانی توی تصمیم گیری داره ایرادش اینجا بود که ممکن بود مدلمون ضعیف باشه مثلsvm که توی تصمیم گیری سهم یکسانی با مدل های دیگه که بهتر بودن داشت خوب ما می خوایم اگر یک مدل ضعیف هست سهم کمتری بگیره نسبت به سایر مدل هامون به این نحو که این مدل دو سطح داره سطح اول که میاد یک سری مدل ها رو آموزش میده و کنار هم با هم ترکیب میشن و بر اساس داده آموزشیمون خروجی داده جدید رو تخمین میزنن،، توی تکنیک ویتینگ ما یک رای گیری ساده داشتیم میامدیم خروجی اینا رو رای گیری میکردیم یا اگر بحث رگراسیون بود میامدیم یک میانگین میگرفتیم و کار تموم بود جواب رو بر اساس بیشترین رای می دادیم

تو سطح دوم تکنیک وتینگ یک مدل دیگه داریم که میاد بر اساس یادگیری عمیق به هر کدوم از جواب ها ی مدلهامون یک وزن میده که این وزن درجه اهمیت مدل رو برای رای گیری مشخص می کنه مثلا توی مثال جلسه پیش گفتیم ما باید یک تکنیک داشته باشیم که درصد ریسک بیمار رو مشخص کنه که اگر مثلا درصد سرطانی بودنش 40 به 60 بود 40 درصد رو مهمتر از 60 درصد سالم بودن بیمار در نظر بگیره و بیمار رو برای آزمایشات تکمیلی بستری کنه.... ما توی این مدل هم بر اساس سطح دوم میایم میگیم حتی اگر یک مدل پایمون مثل svm خوب عمل نکرد بیاد اصلا وزن منفی صفر بگیره توی این مدل هر کدوم از مدل هامون که وزنش بیشتر باشه توی رای گیری سهم بیشتری داره. پس اینم شد مدل استکنیک بریم سراغ اسلاید بعد که مزایای تکنیک استکینگ هست.

اسلاید 15

توانایی افزایش دقت:

 ۱- با ترکیب چندین مدل، استکینگ معمولاً به عملکرد پیش‌بینی بهتری نسبت به مدل تنها منجر می‌شود.

انعطاف پذیری:

 ۲- اجازه می‌دهد از مدل‌های پایه متنوع و انتخاب انعطاف‌پذیر META MODEL استفاده شود.

کاهش اورفیت:

 ۳- با ترکیب مدل‌هایی که به طور جداگانه ممکن است اور فیت داشته باشند، به کاهش اورفیتینگ کمک می‌کند.

اسلاید 16

یادگیری ماشین با استفاده از روش Bagging خب مختصری از دو مدل بعدی رو توضیح میدم تا ارائه زیاد طولانی نشه توی مدل بگینگ ما دوتا اصطلاح داریم یکی bootstrap که به کمک این مفهوم میایم دیتاست آموزشیمون رو به چند قسمت تقسیم میکنیم تا مدل ها فقط با دیتای خودشون آموزش ببینن توی مدل بگینگ باید تمامی مدل های آموزشی یکسان باشن ولی قوی مثلا باید تمامی مدل ها کا ان ان باشن یا درختی باشن نمی تونیم چنتا مدل رو با هم ترکیب کنیم ببنید این مدل برای زمانی خوبه که دیتا ستمون خیلی بزرگ باشه یکی از مشکلاتی که مدل ها دارن اینه که دیتا ست آموزشیمون اگر خیلی بزرگ باشه نمی تونن به درستی آموزش ببینن توی روش بگینگ اومده این مشکل رو برطرف کرده میاد با استفاده از boot strap دیتاست بزرگمون رو به چند قسمت تقسیم میکنه و یک دیتا ست جدید برای هر مدل میسازه و هر قسمت رو در اختیار یک مدل قرار میده تا بتونه آموزش ببینه اینطوری مدل ها تونستن تمام فضاهای ویژگی رو هر چقدر هم زیاد باشه بشناسن پس در مقابل روش های تکی پرفورمنس به شدت بالا میره مثلا به جای چند مدل کا ان ان یک مدل کا ان ان با دیتا ست حجیم داشته باشین هرچقدر هم این مدل کا ان ان خوب باشه ممکنه دچار اور فیتینگ بشه. و با استفاده از تکنیک aggregation میاد این مدل ها رو ترکیب میکنه و در پایان با روش مکس ویتینگ تجمیع انجام میشه و به نتیجه میرسه.

اسلاید 17

میریم سراغ مزایای تکنیک بگینگ در یادگیری جمعی

کاهش اورفیت:

۱- با آموزش چندین مدل بر روی زیرمجموعه‌های مختلف داده، بگینگ خطر اورفیت که یک مدل تکی ممکن است بر روی داده‌های آموزش اصلی داشته باشد، را کاهش می‌دهد.

بهبود دقت:

 ۲- جمع‌آوری پیش‌بینی‌ها از چندین مدل به عملکرد بهتری نسبت به یک مدل تکی منجر می‌شود، زیرا از قدرت‌های چندین فرضیه بهره می‌برد.

پایداری مدل:

 ۳- بگینگ پایداری و استحکام مدل را افزایش می‌دهد، با کمینه کردن تأثیر نقاط داده‌ای نویزی و نقاط پرت.

اسلاید 18

یادگیری ماشین با استفاده از روش Boosting

خب مدل بوستینگ باعث قدرت گرفتن یک مدل ضعیف مثل svm میشه مثلا شما در نظر بگیرین که یک مسئله دارین با سه مدل داده یا همون سه دیتاست و می خواین با روش svm این سه داده رو از هم جدا کنید svm برای مسائل دو کلاسه ممکنه خوب عمل کنه ولی برای مسائل غیر خطی به شدت بد عمل میکنه توی روش بوستینگ ما میایم میگیم شما به جای استفاده از یک مدل ضعیف مثل svm بیاد چنتا svm رو کنار هم قرار بده ومسئله ما رو که مسئله چند کلاسه هست حل کنه مثلا توی این شکل شما در نظر بگیرید اگر من بخوام فقط با استفاده از اس وی ام این داده ها رو از هم جدا کنم می تونم دوتا از داده ها رو از هم تشخیص بدم و بینشون خط بکشم اما داده های سومم مشخص نیست که باید کجا قرار بگیره حالا اگر از سه تا اس وی ام استفاده کنم برای حل این مسئله کافیه بیام بگم مثلا دیتا ست اول و دوم رو به اس وی ام بده تا بینشون خط رسم کنه و به اس وی ام دوم بگم دیتا ست اولم رو با سوم جدا کن و به اس وی ام سوم بگم دیتا ست یا همون دیتاهای دومم رو با سومم جدا کن به این ترتیب ما مواجه با چنین شکل فرضی میشیم که میبینید تونسته هر سه مدل داده ی ما رو به صورت درست تخمین بزنه . پس این شد روش بوستینیگ ..

توی روش بوستینگ ما میایم چند مدل ضعیف رو به عنوان مدل های پایمون در نظر میگیریم توجه داشته باشید که مدل های پایه همه باید ضعیف باشن و همنوع مثلا همه باید اس وی ام باشن نه چیز دیگه بعد میاد مرحله بعد تمامی دیتا ها رو به مدل اولمون تحویل میده این مدل تا جایی که میتونه داده ها رو تخمین میزنه مثلا 20 درصد درست تخمین زده و برای 80 درصد بعدی نتونسته تخمینی بزنه اون 80 درصدی که نتونسته میشن دیتاست ورودی اس وی ام دوم ما باز اس وی ام دوم میاد تا یک حدی روی این دیتا تخمین درست میزنه مثلا 20 درصد هم اس وی ام دوم تخمین درست میزنه و اون 60 درصد باقی مانده میشه ورودی مدل سوممون یا همون اس وی ام سوم این روال تا جایی تکرار میشه که بتونن کل فضای دادها رو درست تخمین بزنن. میریم سراغ مزایای بوستینگ

اسلاید 20

بهبود عملکرد:

۱- با تمرکز بر روی اشتباهات مدل‌های قبلی، بوستینگ معمولاً به طرز قابل توجهی عملکرد کلی را بهبود می‌بخشد.

انطباق‌پذیری:

 ۲- بوستینگ می‌تواند با تنظیم تمرکز بر روی نمونه‌های دشوار، به پیچیدگی داده‌ها سازگار شود.

کاهش بایاس:

 ۳- با تمرکز تکراری بر روی اشتباهات، بوستینگ بایاس مدل کلی را کاهش می‌دهد.

اسلاید 22

در یادگیری ماشین، کاهش خطا و واریانس مدل‌ها یکی از عوامل کلیدی تعیین‌کننده موفقیت فرایند یادگیریه. در مطالعات مختلف که این مقاله روشون تمرکز داشته، اثبات شده که ترکیب خروجی الگوریتم‌های مختلف طبقه‌بندی می‌تونه خطای تعمیم رو بدون افزایش واریانس مدل کاهش بده. که این نکته اصلی جوهره‌ای به نام یادگیری جمعی یا ترکیبیه . تلاش‌های پژوهشی فراوانی یادگیری ترکیبی رو بر یادگیری تک‌مدلی در حوزه‌های مختلف ترجیح دادند. مزیت اصلی یادگیری ترکیبی، ترکیب چندین مدل مستقل برای بهبود عملکرد پیش‌بینی و ایجاد یک مدل قوی‌تر بوده که از تک‌تک مدل‌ها برتر عمل می‌کنه. در ادبیات، تکنیک‌های مختلفی برای تقویت الگوریتم‌های طبقه‌بندی وجود داره که تفاوت اصلی بین هر دو روش ترکیبی در آموزش مدل‌های پایه و نحوه ترکیبشون با همدیگست.

همچنین چندین تلاش پژوهشی یادگیری ترکیبی رو به مدل‌های یادگیری عمیق معرفی کرده‌اند تا مشکلاتی را که در طول فرایند یادگیری مدل‌های یادگیری عمیق به وجود می‌آمدند، برطرف کنن. معمولاً چالش اصلی مدل‌های یادگیری عمیق این بود که نیاز به دانش و تجربه زیادی برای تنظیم بهینه پارامترهای به منظور دستیابی به حداقل خطا رو داشتند. با این حال، یافتن پارامترهای بهینه نیازمند تکنیکی خسته‌کننده در فضای جستجوی ماست که به نوبه خودش به وظیفه‌ای زمان‌بر و پرهزینه تبدیل شده. بنابراین، چندین تلاش پژوهشی یادگیری عمیق ترکیبی رو در بسیاری از حوزه‌ها به کار برده‌اند و بیشتر این تلاش‌ها حول روش‌های ترکیبی ساده متمرکز شده‌اند.

در پایان توی این مقاله ما اومدیم چند استراتژی از یادگیری جمعی که Ensemble learning نام داشت مثل روش وتینگ ، استکینگ ، بگینگ و بوستینگ رو مورد بررسی قرار دادیم و با نحوه عملکرد هر کدوم از این روش ها برای استفاده و ترکیب با یادگیری عمیق آشنا شدیم ارائه من تمام شد ممنون از توجه شما.