



الگوریتم‌های بکار رفته شامل روش‌های مختلف یادگیری ماشین هستند. بالاترین سطح زیر منحنی راک (AUC-ROC) مربوط به RF (۰/۹۷۵) بود، و XGBoost و Gradient Boosting به ترتیب با مقادیر ۰/۹۷۳ و ۰/۹۷۲ در رتبه‌های بعدی قرار گرفتند. در نهایت، بهبود دقت و کارایی مدل‌ها با استفاده از تکنیک‌های پیشرفته یادگیری عمیق و بهینه‌سازی پارامترها پیشنهاد شد (۱۱).

Sriniva و همکاران یک مدل هیبریدی- نزدیکترین همسایگان (CNN-KNN) را ارائه کردند که به طبقه‌بندی تومور مغزی با استفاده از تصاویر MRI با دقت بالاتر و نرخ خطای کمتر کمک می‌کند. آنها شبکه‌های عصبی کانولوشن را برای استخراج ویژگی‌های k- نزدیکترین همسایه جهت طبقه‌بندی بکار گرفتند. بر اساس اطلاعات موجود، بالاترین دقت و اف ۱ اسکور مدل CNN-KNN برابر با ۹۶/۲۵ درصد برآورد شد. همچنین، نرخ خطای (error rate) آن مدل ۷۵/۳ درصد تخمین زده شد (۱۴).

تحقیق بر روی بهینه‌سازی بیشتر مدل‌های هیبریدی و بررسی استفاده از دیگر الگوریتم‌های یادگیری عمیق و ترکیب آن‌ها با KNN برای بهبود دقت و کارایی به محققان علاقه‌مند به کار در آن حوزه پیشنهاد شد (۱۴).

در مطالعه‌ای دیگر رویکردهای یادگیری ماشین برای پیش‌بینی تومورهای مغزی با استفاده از مدل‌های ماشین رگرسیون لجستیک (LR)، KNN، SVM، SVM Kernel، بیز ساده (NB)، DT، جنگل تصادفی (RF) و Cat Boost بررسی و مقایسه شدند. بالاترین دقت، صحت و امتیاز F1 ۰/۹۸ و بالاترین یادآوری ۰/۹۹ محاسبه شد. در نهایت، کارهای پیشنهادی برای آینده شامل بهبود دقت مدل‌ها و استفاده از داده‌های بیشتر برای آموزش بهتر الگوریتم‌ها است (۱۵).

مطالعه‌ی Wang و همکاران با هدف شناسایی روش‌های یادگیری ماشینی بهینه برای درجه‌بندی دقیق گلیوم با استفاده از MRI بود. در این مطالعه از روش‌های یادگیری ماشین شامل ماشین بردار پشتیبان-حذف ویژگی بازگشتی (SVM-RFE)، حداقل انقباض مطلق و عملگر انتخاب (LASSO)، حداکثر ارتباط و حداقل افزونگی (mRMR)، DT، KNN، NB، LR استفاده شد. مدل‌های SVM-RFE و SVM به معیارهای دقت ۹۱/۱ درصد، حساسیت ۹۴/۲ درصد، ویژگی ۸۹ درصد و مساحت زیر منحنی (AUC) ۰/۹۸۵ دست یافتند. آن مطالعه نشان داد که الگوریتم SVM-RFE همراه با سایر روش‌های یادگیری ماشینی، می‌تواند ابزاری غیر تهاجمی قابل قبولی برای درجه‌بندی گلیوم پیش از جراحی فراهم کند و مدیریت بیمار را تسهیل کند. توصیه‌های آینده شامل اعتبار سنجی بیشتر مدل‌های پیش‌بینی در جمعیت‌های بزرگتر و متنوع‌تر بیماران است (۱۶). در تمامی این مطالعات، پژوهشگران از تکنیک‌ها و مجموعه

کاربردی کلینیکی ایران و بسیاری از کشورهای جهان تشخیص تومور در تصاویر MRI بصورت دستی انجام می‌گیرد که بسیار وقت‌گیر است (۸). ارزیابی و پردازش تصاویر پزشکی در بالین از اهمیت ویژه‌ای برخوردار است و به یکی از سودمندترین راهبردها برای تشخیص و شناسایی بیماری‌های مختلف تبدیل شده است. تجسم و رمزگشایی عکس‌ها به پزشکان و رادیولوژیست‌ها اجازه‌ی تجزیه و تحلیل اختصاصی را می‌دهد (۹) و به همراه اطلاعات بالینی می‌تواند تأثیر بالایی در تشخیص زودهنگام داشته باشد (۱۰). کاربرد هوش مصنوعی و یادگیری ماشین (ML) بصورت روزافزون در حال گسترش است (۱۱) تا با تکیه بر پیشرفت‌های آن حوزه، تومورهای مغزی با دقت بالاتری تشخیص داده شوند و تصمیم‌های آگاهانه در خصوص برنامه‌های درمانی بیماران اتخاذ نمایند. با توجه به آنکه ML توانایی شناسایی، تشخیص و درمان انواع مختلف تومورهای مغزی را دارد، زمینه مطالعاتی نسبتاً جدیدی باز شده و با گذشت زمان توجه بسیاری را به خود جلب کرده است. یادگیری ماشین، داده‌ها را برای پیش‌بینی نتایج و قضاوت تحلیل می‌کند. از این رو به نظر می‌رسد استفاده از الگوریتم‌ها و رویکردهای متنوع برای شناسایی و دسته‌بندی انواع مختلف تومور مغزی با استفاده از ML ضروری است (۱۲). در سال‌های اخیر پژوهش‌های متعددی در زمینه استفاده از الگوریتم‌های یادگیری ماشین برای تشخیص تومورهای مغزی بر اساس تصاویر ام‌آرآی انجام شده است. برخی از کارهای مربوط به آن حوزه به اختصار در ادامه‌ی مطالب بحث شده است.

Praveena و همکاران مطالعه‌ای را با هدف بررسی یک رویکرد یادگیری ماشین برای تقسیم‌بندی، ارزیابی و پیش‌بینی رشد تومور مغزی با استفاده از تصاویر MRI طراحی کردند (۱۳). دو مدل یادگیری ماشین شامل ماشین بردار پشتیبان (SVM) و درخت تصمیم (DT) استفاده شد. الگوریتم بکار رفته شامل خوشه‌بندی K-means برای تقسیم‌بندی و ماتریس هم‌رخدادی سطح خاکستری (GLCM) برای استخراج ویژگی‌ها بود. بالاترین صحت (Precision) ۹۵/۹۱ درصد است که توسط SVM به دست آمد. دقت (Accuracy) برای هر دو مدل SVM و DT برابر با ۹۰ درصد بود. یادآوری (Recall) برای SVM برابر با ۹۰/۳۸ درصد و برای DT برابر با ۸۸/۴۶ درصد است. همچنین، اف ۱ اسکور (F1 Score) در SVM بیشتر از DT برآورد شد. نتایج آنان بررسی طبقه‌بندی‌های مختلف و روش‌های بهینه‌سازی برای بهبود عملکرد را برای آینده پیشنهاد داد (۱۳). به منظور مقایسه الگوریتم‌های یادگیری ماشین برای تشخیص و طبقه‌بندی تومور مغزی مدل‌های ماشینی SVM، KNN، جنگل تصادفی (RF) و XGBoost تجزیه و تحلیل شدند.

عصبی، باید از یکسان بودن اندازه‌ی تصویر ورودی اطمینان حاصل کرد، زیرا استفاده از شبکه‌ی عصبی کانولوشن برای آموزش یک شبکه‌ی عصبی آسان نیست. بر همین اساس تصاویر با استفاده از کدنویسی پایتون هم‌اندازه شدند (۲۰۰×۲۰۰). توزیع و مراحل داده‌های ورودی برای این مدل‌ها به طور خلاصه در شکل ۳ نشان داده شده است.

#### مدل‌های یادگیری ماشین

هدف اصلی یادگیری ماشین، طبقه‌بندی داده‌ها است. چنین رده‌بندی‌هایی ممکن است باینری (دودویی) یا چندگانه باشد که در این صورت از طبقه‌بندی‌های باینری مانند درخت تصمیم، K-نزدیک‌ترین همسایه، ماشین بردار پشتیبان، رگرسیون لجستیک و جنگل تصادفی و بیز ساده استفاده می‌شود.

رگرسیون لجستیک یک نوع الگوریتم یادگیری ماشینی نظارت شده است که برای اهداف طبقه‌بندی یا پیش‌بینی نتایج دودویی کاربرد دارد (۱۷) و دارای یک روش آماری است که برای درک و طبقه‌بندی رویدادهای پیچیده و فازی از آن استفاده می‌کند (۱۸). درخت تصمیم (Decision Tree)، نوعی الگوریتم یادگیری ماشینی غیرپارامتریک و نظارت شده است که در طبقه‌بندی و رگرسیون استفاده دارد (۱۷).

جنگل تصادفی (Random Forest)، این الگوریتم یادگیری ماشینی نظارت شده برای کارهای طبقه‌بندی و رگرسیون بکار می‌رود. بر اساس RF (۱۹) به جای استفاده از بهترین شاخه در بین همه‌ی متغیرها، گره اولیه با استفاده از بهترین متغیرهای تصادفی انتخاب شده و توسط شاخه‌ها به گره‌های بعدی متصل می‌گردد. بیز ساده (Naïve Bayes)، یکی از انواع الگوریتم یادگیری ماشینی نظارت شده است. روش مذکور برای کارهای طبقه‌بندی بر اساس قضیه بیز کاربرد دارد و فرض می‌شود ویژگی‌های داده‌ها مستقل از یکدیگر هستند (۱۷). ماشین بردار پشتیبان (Support Vector Machine) الگوریتمی است که به دلیل دقت بالا می‌تواند به راحتی در داده‌های دشوار و پیچیده استفاده شود. k-نزدیکترین همسایه (K-Nearest Neighbor)، یک تکنیک طبقه‌بندی ساده و قوی است که در آن بردار ویژگی تست با یافتن k نزدیک‌ترین بردار همسایه آموزشی طبقه‌بندی می‌شود. فاصله‌ی بین بردار آموزش و آزمون با تکنیک اندازه‌گیری متریک فاصله متفاوت محاسبه می‌گردد (۱۹).

#### آموزش و آزمایش الگوریتم‌ها

مجموعه داده‌ی این پروژه شامل یک پوشه حاوی ۲۴۰۰ تصویر برچسب دار می‌باشد. به منظور آماده‌سازی داده‌ها برای آموزش و آزمایش، مجموعه داده با استفاده از کتابخانه Scikit-Learn و تابع train\_test\_split به دو مجموعه تقسیم شد. تقسیم‌بندی به ترتیب بر روی ۸۰ و ۲۰ درصد تصاویر در مجموعه تمرین و تست انجام

داده‌های متفاوتی استفاده کرده‌اند. برخی از داده‌های سازمان‌های مراقبت‌های بهداشتی استفاده کرده‌اند؛ درحالی‌که برخی داده‌ها را از مخازن آنلاین جمع‌آوری کرده‌اند. کارایی MRI و ML در تشخیص و طبقه‌بندی بیماری‌های مختلف در مطالعات قبلی ارزیابی شده است. از آنجایی که تشخیص تومورهای مغزی با دست‌زمان بر و همراه با خطا می‌باشد و یادگیری عمیق هم با توجه به اهمیت و پیچیدگی آن در تشخیص تومورهای مغزی و اثر آن بر سلامت افراد کمی زمان‌بر است. لذا هدف از مطالعه‌ی حاضر، استفاده از تصاویر MRI، الگوریتم‌های یادگیری ماشین که شامل SVM، LR، KNN، NB، DT، RF برای تشخیص تصاویر توموری و غیر توموری در کمترین زمان لازم می‌باشد و تلاش می‌شود بهترین مشورت گروهی توسط الگوریتم‌ها برای تشخیص دقیق تصاویر توموری و غیرتوموری روی مجموعه داده‌ها گرفته شود.

#### روش‌ها

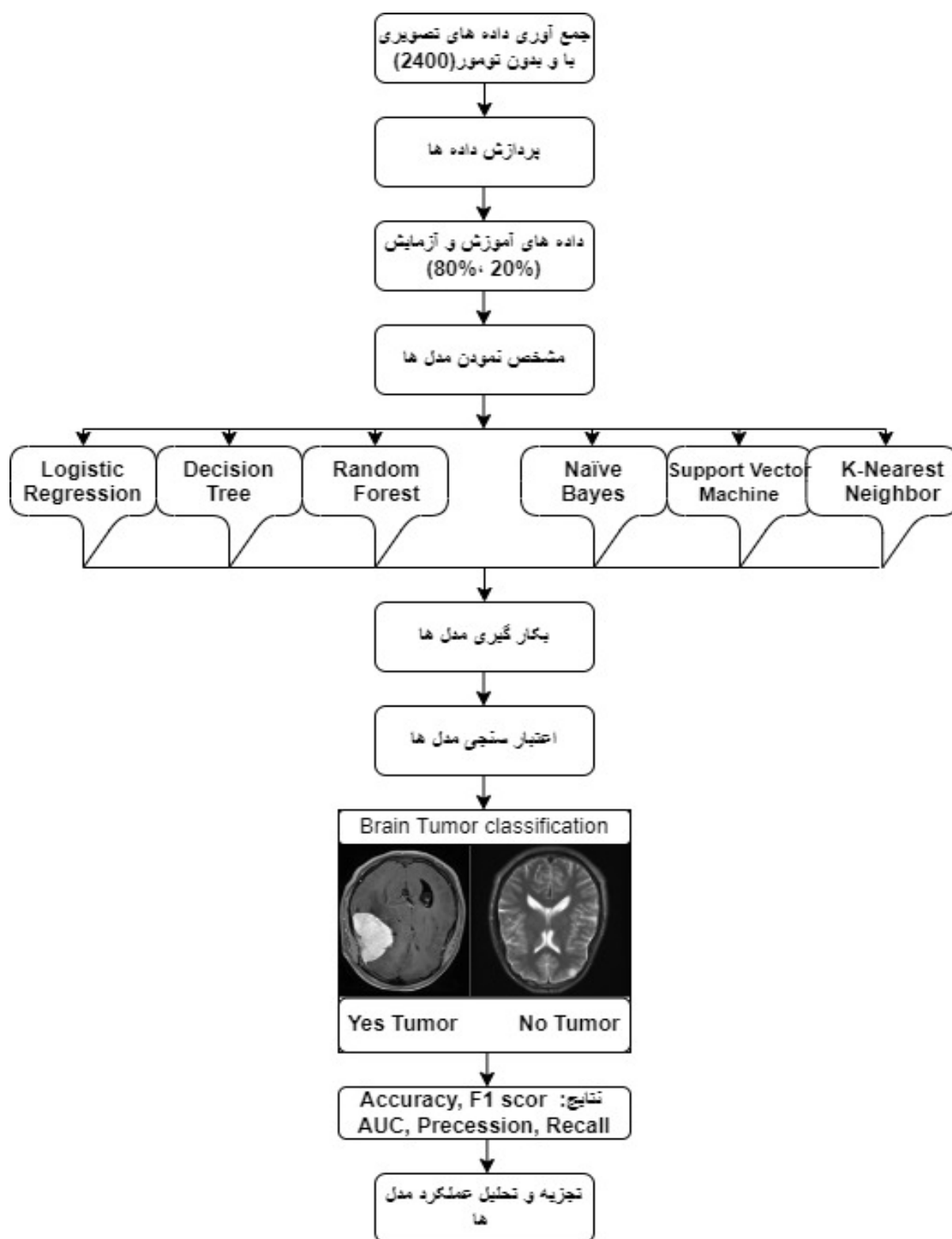
##### راه‌اندازی تجربی، فراخوانی کتابخانه‌ها

تمرکز اصلی مطالعه‌ی حاضر بر ارزیابی مدل‌های یادگیری ماشین برای شناسایی تصاویر توموری و غیر توموری (سالم) با استفاده از تصاویر MRI می‌باشد تا بتوان مدل‌های برتر را معرفی نمود. تعریف و پیاده‌سازی مدل با یک پلتفرم کد نویسی با ویژگی‌های زیر انجام شد: سیستم عامل (ویندوز)، محیط اجرای کد (CPU: i7.4710k، GPU: Nvidia GT740-2G، G D3۱۶، H.D.D: 1T، PC، رم: ۸GB، S.S.D: 240G)، زبان برنامه‌نویسی (Python)، چارچوب نرم‌افزار (TensorFlow)، کتابخانه‌ی سطح بالا (NumPy، Matplotlib، Pandas). شکل ۱ روش پیشنهادی را نمایش می‌دهد.

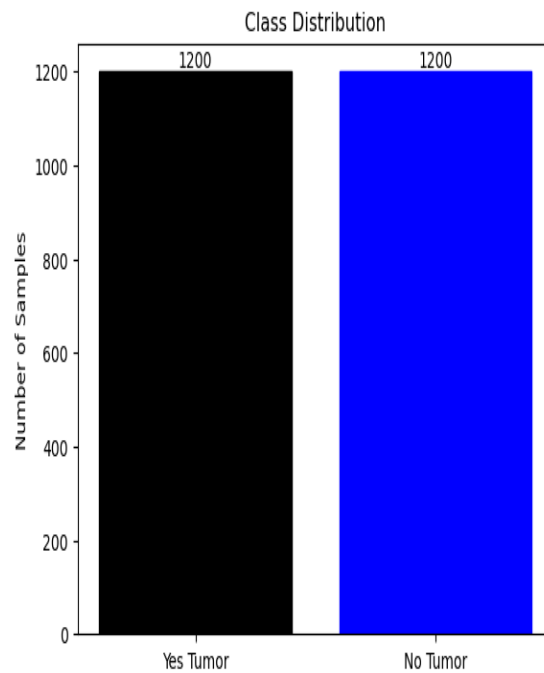
##### جمع‌آوری تصاویر MRI مغزی

مجموعه تصاویر MRI از Kaggle.com جمع‌آوری شده است که یکی از عمومی‌ترین پایگاه داده‌های پزشکی برای آزمایش الگوریتم‌هایی مانند طبقه‌بندی، تشخیص تومورها و تقسیم‌بندی تصاویر پزشکی است. به منظور انجام طبقه‌بندی، ۲۴۰۰ تصویر MRI با کیفیت با فرمت JPG از وب‌سایت مذکور جمع‌آوری شد. کلاس «بله» شامل ۱۲۰۰ تصویر از تومورهای مغزی و کلاس «نه» شامل ۱۲۰۰ تصویر از مغزهای سالم است (شکل ۲).

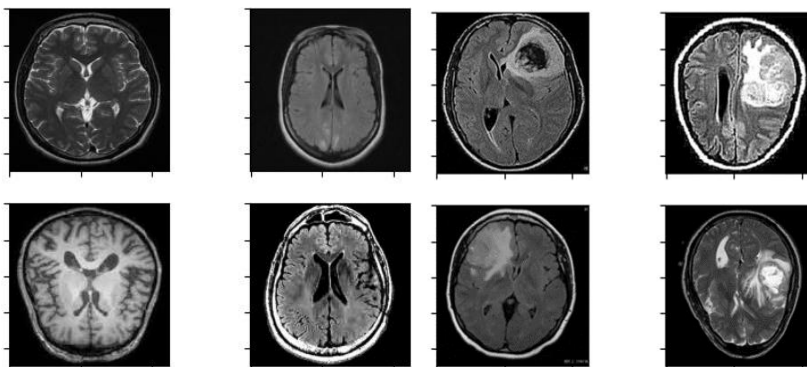
قبل از ورود به مرحله‌ی آموزش شبکه‌ها، باید پیش‌پردازش روی مجموعه‌ی داده انجام شود تا داده‌ها بتوانند انتظارات و شرایط لازم را برای ورود به آموزش را پیدا کنند و با ورودی شبکه‌ی ساختار برای آموزش و ارزیابی مدل مطابقت داشته باشند. تصاویر جمع‌آوری شده اندازه‌های متفاوتی دارند (تفاوت در طول و عرض، به عنوان مثال، ۱۶۸ × ۳۰۰ یا ۲۹۰ × ۲۵۰). برای استفاده از این تصاویر در شبکه‌ی



شکل ۱. پیاده‌سازی مدل با یک پلتفرم کد نویسی



شکل ۲: توزیع مجموعه تصاویر با و بدون تومور



شکل ۳: نمایش تصاویر با اندازه‌ی یکسان در قالب با و بدون تومور با اندازه‌ی ۲۰۰×۲۰۰×۳

تصاویر MRI بر اساس مطالعه و اجرای طبقه‌بندی تومور استفاده شد. پارمترهایی از قبیل  $C=0.1$  برای الگوریتم رگرسیون لجستیک، برای الگوریتم  $K$ -نزدیک‌ترین همسایه و  $n\_estimator=100$  جهت جنگل تصادفی بکار گرفته شد.

**ارزیابی عملکرد الگوریتم‌ها (تنظیم ماتریس سردرگمی و منحنی ROC)**  
این بخش اطلاعات مربوط به معیارهای ارزیابی عملکرد مورد استفاده در این مطالعه را ارائه می‌دهد. از معیارهای دقت، صحت، یادآوری و امتیاز F1 برای ارزیابی عملکرد مدل‌ها استفاده می‌شود. معیارهای مذکور از مثبت واقعی (TP)، منفی واقعی (TN)، مثبت کاذب (FP) و منفی کاذب (FN) بدست می‌آیند. بنابراین علاوه بر ROC، ماتریس سردرگمی هم می‌بایست ارزیابی گردد.

شد. دو پوشه "train" و "test" ایجاد شد که هر کدام شامل یک کلاس ۱۲۰۰ تصویر بود. عمل نرمالیزه کردن داده‌ها با استفاده از  $preprocessing.MinMaxScaler=scaler()$  صورت پذیرفت. بدین معنا که داده‌ها به فرمتی تبدیل شدند که تمامی ویژگی‌ها در یک مقیاس خاص قرار داشته باشند. با توجه به آنکه بسیاری از الگوریتم‌ها به مقیاس ویژگی‌ها حساس هستند، چنین فرایندی معمولاً برای بهبود عملکرد الگوریتم‌های یادگیری ماشینی انجام می‌شود. به طور کلی،  $MinMaxScaler$  نقش مهمی در پیش‌پردازش داده‌ها دارد و می‌تواند به بهبود دقت الگوریتم‌های یادگیری ماشینی در پیش‌بینی کردن خروجی‌ها کمک کند. از شش الگوریتم یادگیری ماشینی ذکر شده برای پیش‌بینی نوع تصاویر توموری و بدون توموری با استفاده از

شامل LR، DT، RF، NB، SVM و KNN می‌باشند. هر یک از این الگوریتم‌ها بر اساس پارادایم‌های خاصی عمل می‌کنند و قابلیت‌های منحصر به فردی برای تجزیه و تحلیل و تفسیر داده‌های تصویر دارند. این الگوریتم‌ها با استفاده از مجموعه‌ای از تصاویر MRI آموزش داده شد و دقت آن‌ها در تشخیص تصاویر ندیده آزمایش شد. برای ارائه‌ی ارزیابی دقیق‌تر از عملکرد الگوریتم‌ها، نتایج آموزش و آزمایش، منحنی‌های ROC، ماتریس سردرگمی و پیش‌بینی ۳۰ تصویر از بخش تست بر اساس مدل‌ها مورد بررسی قرار گرفت. این تحلیل‌ها در ارزیابی عملکرد هر الگوریتم در جنبه‌های مختلف مانند دقت، حساسیت، ویژگی و انتخاب بهترین مدل برای تشخیص تومورهای مغزی کمک می‌کند.

#### نتایج آموزشی و آزمایشی مدل

در ارزیابی عملکرد مدل، دو مورد از این پارامترها شامل دقت و خطا (از دست دادن) است. دقت نشان می‌دهد که چند درصد از پیش‌بینی‌های مدل صحیح است. خطا معیار معکوس دقت است. علاوه بر این، داده‌های اعتبارسنجی در طول آموزش تا حدی برای ارزیابی کیفیت مدل آموزش دیده استفاده می‌شود. این داده‌های اعتبارسنجی در هر قسمت آموزشی گنجانده شده است تا از برازش یا عدم تناسب مدل ساخته شده با توجه به داده‌های تست جلوگیری شود. شکل ۴ مقادیر آموزش و آزمایش این الگوریتم‌ها را نشان می‌دهد.

AUC-ROC یک معیار کلی از عملکرد در تمام آستانه‌های طبقه‌بندی‌های ممکن را ارائه می‌دهد (۲۰)، در حالیکه CA به این نکته اشاره می‌کند که چقدر پیش‌بینی دقیق است و به صورت زیر محاسبه می‌گردد:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \quad (1)$$

از طرفی صحت و یادآوری دو معیار مهم برای ارزیابی عملکرد مدل‌های یادگیری ماشین، به خصوص در مسائل با کلاس‌های نامتعادل، هستند. این دو معیار به ما کمک می‌کنند تا عملکرد پیش‌بینی کننده را در شناسایی صحیح عناصر مختلف در داده‌ها ارزیابی کنیم (۲۱).

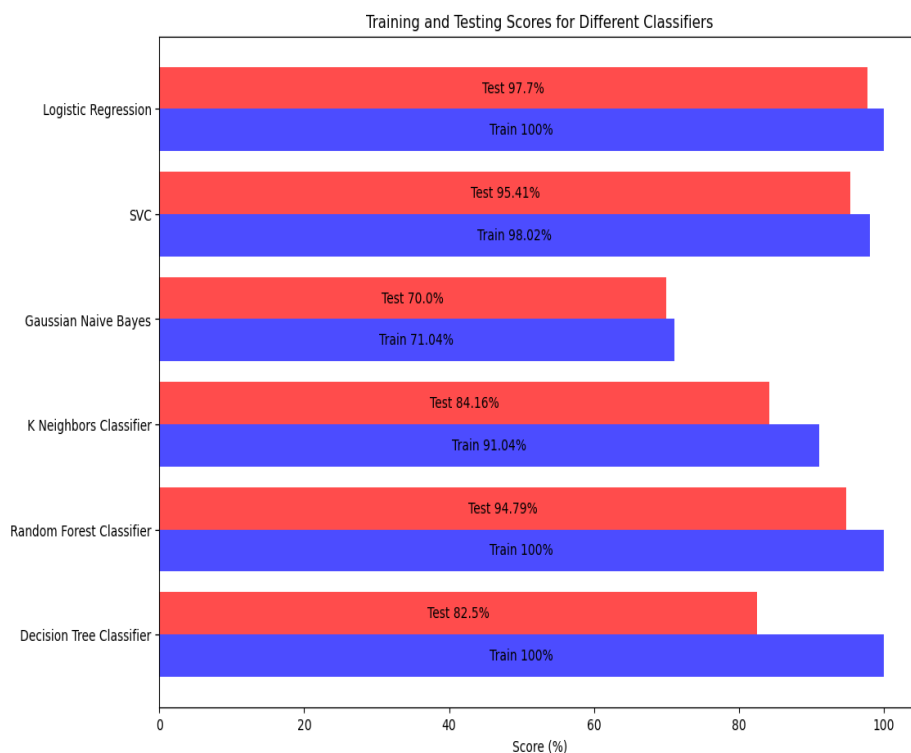
$$\text{Precision} = \frac{TP}{TP+FP} \quad (2) \quad \text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

امتیاز F1، محاسبه متریکی دیگری است که بر اساس میانگین هارمونیک صحت و یادآوری است، بنابراین بهبود همزمان در صحت و یادآوری منجر به امتیاز بالاتر F1 می‌شود (۲۲). بدیهی است که این معیار به ویژه در شرایطی که نیاز به ارزیابی جامع عملکرد مدل بر اساس صحت و یادآوری وجود دارد، ارزشمند است.

$$\text{F1score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

#### یافته‌ها

در این مقاله، ما از ۶ الگوریتم یادگیری ماشینی مختلف برای تشخیص تومورهای مغزی در تصاویر MRI استفاده کردیم. این الگوریتم‌ها



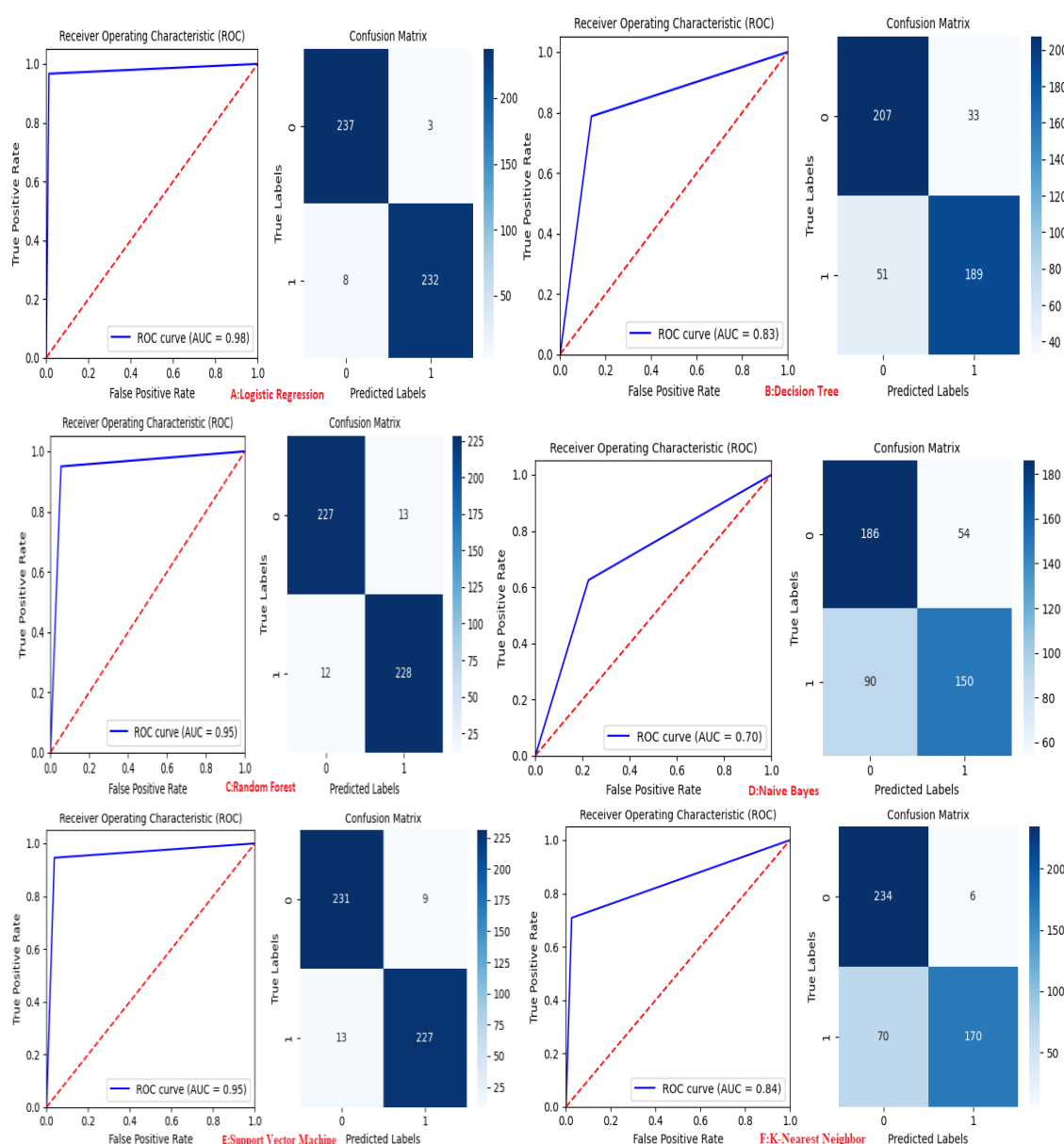
شکل ۴: نمودار نتایج آموزش و آزمون الگوریتم‌ها

## بحث

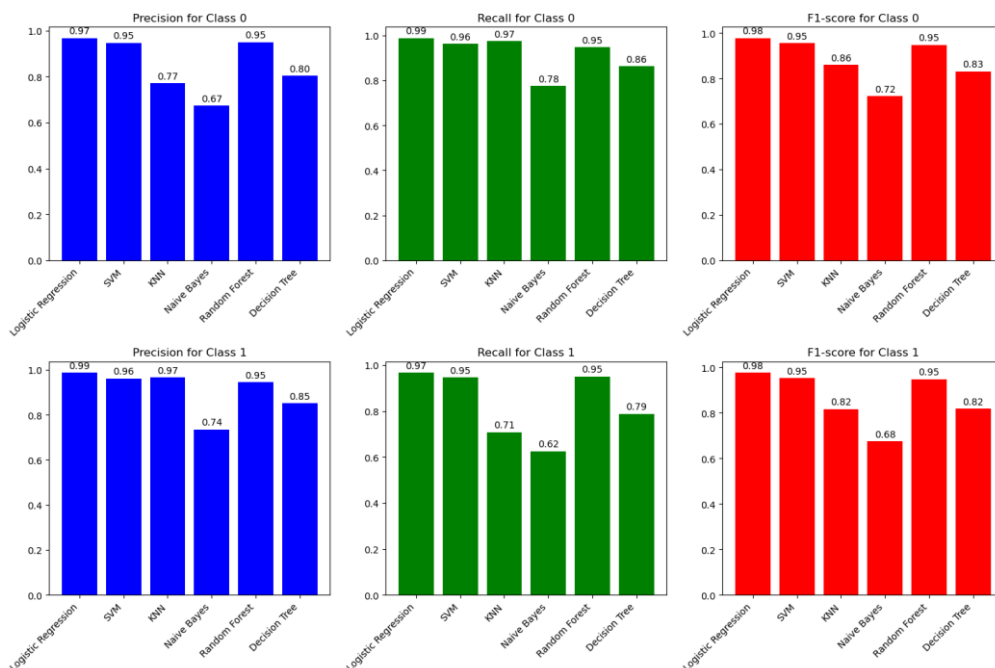
بر اساس شکل ۴، برای ارزیابی دقیق الگوریتم‌های یادگیری ماشین، دو مرحله‌ی کلیدی آموزش و آزمایش وجود دارد: در مرحله‌ی آموزش، الگوریتم از داده‌ها یاد می‌گیرد، در حالی که در مرحله‌ی آزمایش، توانایی تعمیم‌پذیری الگوریتم به داده‌های جدید و ندیده مورد ارزیابی قرار می‌گیرد. الگوریتم رگرسیون لجستیک دقت آموزشی کامل (۱۰۰ درصد) و دقت آزمایشی بالا (۹۷/۷ درصد) دارد. الگوریتم ماشین پشتیبان دقت آموزشی بالا (۹۸/۰۲ درصد) و دقت آزمایشی کمی پایین‌تر (۹۵/۴۱ درصد) دارد. این نشان می‌دهد که این دو الگوریتم داده‌های آموزشی را به خوبی یاد گرفته و همچنین توانایی

نتایج منحنی ROC و ماتریس سردرگمی الگوریتم‌ها در پیش‌بینی تصاویر ماتریس‌های سردرگمی و منحنی‌های راک برای شش الگوریتم مطابق شکل‌های A تا F (شکل ۵) می‌باشد و توضیحات لازم در مورد ارزیابی‌ها در قسمت بحث ارائه می‌گردد.

نتایج صحت، بازیابی و  $F1$ -score الگوریتم‌ها و پیش‌بینی نوع تصاویر نشان‌دهنده‌ی عملکرد الگوریتم‌ها در تشخیص و پیش‌بینی نوع تومور و غیرتومور می‌باشد. این معیارها برای ارزیابی دقت و کارایی الگوریتم‌ها در تصمیم‌گیری بر روی داده‌های تصاویر مورد استفاده قرار می‌گیرند (شکل ۶).



شکل ۵: نمودارهای ROC و ماتریس‌های سردرگمی الگوریتم‌ها



شکل ۶: نمایش صحت، فراخوانی و امتیاز FI الگوریتم‌ها

الگوریتم درخت تصمیم با دقت آموزشی کامل (۱۰۰ درصد) دارای دقت آزمایشی (۸۲/۵ درصد) پایین تری در مقایسه با برخی دیگر از الگوریتم‌های دارد. درخت‌های تصمیم می‌توانند در صورتی که خیلی پیچیده شوند، به آموزش داده‌ها بیش برآزش بخورند، که ممکن است در کاهش دقت آزمایشی خود را نشان دهد. در نتیجه، رگرسیون لجستیک و جنگل تصادفی با دقت‌های آزمایشی بالا و توانایی خوب در تعمیم‌پذیری بهترین عملکرد را در این مجموعه داده دارند. ماشین پشتیبان نیز نتایج خوبی را نشان می‌دهد. نایب بیز و درخت تصمیم امکان بهبود دارند، در حالی که عملکرد KNN ممکن است به تنظیمات پارامتری نیاز داشته باشد. در ارزیابی الگوریتم‌های یادگیری ماشین، فقط دقت تنها معیار نمی‌باشد. معیارهای دیگری مانند صحت، بازخوانی، ضریب FI و AUC اطلاعات بیشتری را ارائه می‌دهند، به ویژه برای مجموعه داده‌های نابرابر یا زمانی که هزینه‌های خطاها متفاوت است.

با توجه به داده‌های استخراج شده از شکل ۵، الگوریتم رگرسیون لجستیک با AUC بالای ۰/۹۸ و تعداد بالای TP و TN بهترین عملکرد را دارد و در تشخیص موارد مثبت و منفی بسیار عالی عمل کرده است. الگوریتم درخت تصمیم با AUC کمتر از بقیه الگوریتم‌ها و تعداد زیادی از FP و FN نشان می‌دهد که نیاز به بهبود دارد. جنگل تصادفی با عملکرد خوبی از لحاظ AUC و تعداد TP و TN به خوبی عمل کرده است، در حالی که نایب بیز با AUC پایین‌تر و تعداد زیادی از FP و FN نشان می‌دهد که عملکرد آن نیاز به بهبودی

خوبی در تعمیم به داده‌های جدید دارد. با این حال، دقت آموزشی کامل ممکن است اشاره به *overfitting* داشته باشد، و زمانی رخ می‌دهد که الگوریتم داده‌های آموزشی را به طور مبالغه‌آمیز یاد می‌گیرد (نویز و دیگر عوامل مداخله‌ای). کاهش کمی در دقت آزمایشی ممکن است به این علت باشد که الگوریتم پیچیدگی داده‌ها را به طور کامل نمی‌تواند بهبود بخشد یا کمی *overfitting* داشته باشد.

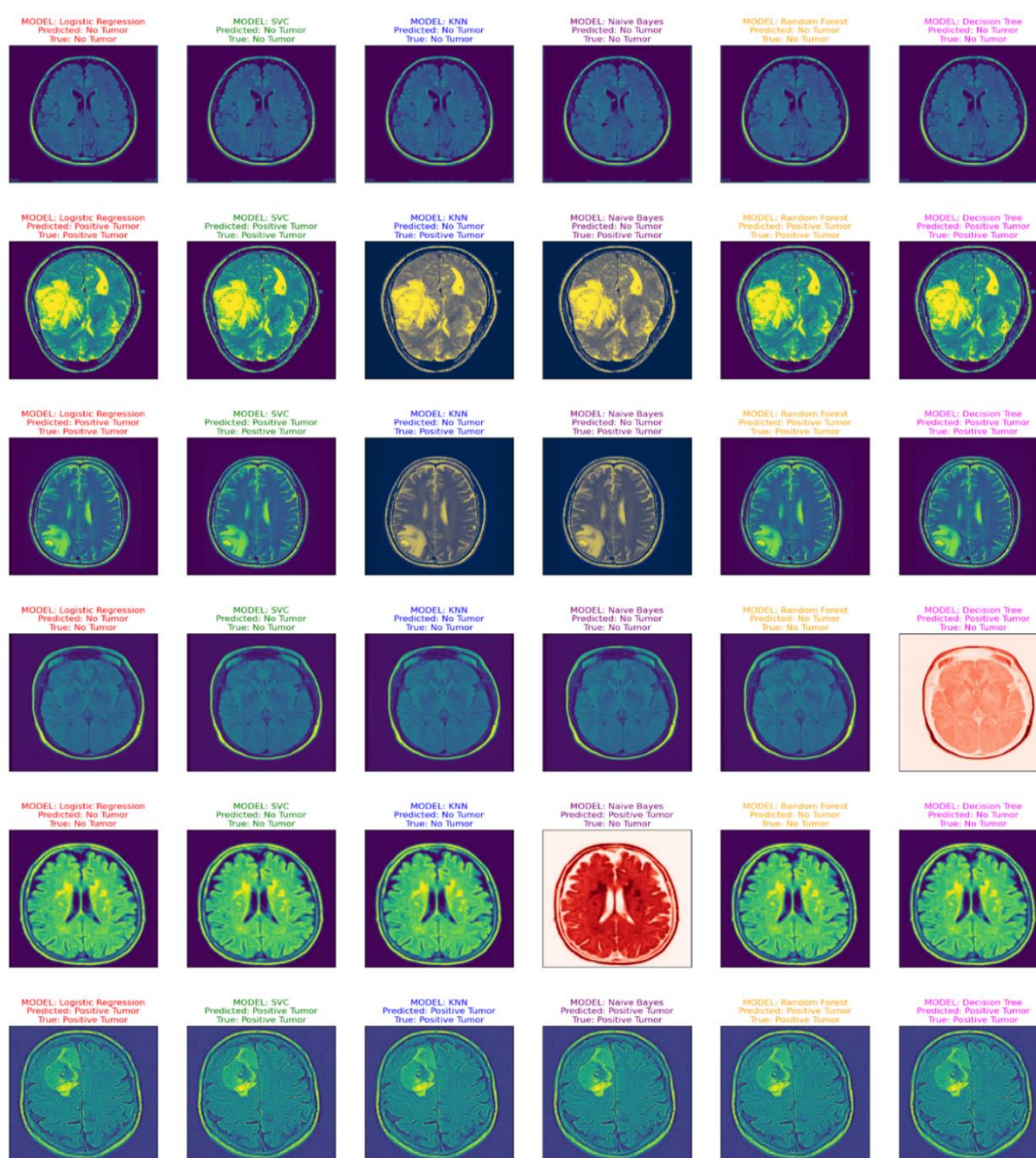
الگوریتم نایب بیز (NB) کمترین عملکرد را در مقایسه با سایر الگوریتم‌ها نشان می‌دهد، با هر دو دقت آموزشی و آزمایشی زیر ۷۵ درصد. این حاکی از آن است که فرضیاتی که توسط الگوریتم نایب بیز اتخاذ می‌شوند (استقلالی بودن ویژگی‌ها) ممکن است برای این مجموعه داده به خوبی صحت نداشته باشند، که منجر به عملکرد ناکافی می‌شود. الگوریتم KNN تفاوت قابل توجهی بین دقت آموزشی (۹۱/۰۴ درصد) و دقت آزمایشی (۸۴/۱۶ درصد) دارد. این ممکن است به این علت باشد که الگوریتم نسبت به انتخاب 'k' (تعداد هم‌سایگان) و ساختار داده‌ها حساس است. کاهش در دقت آزمایشی نشان می‌دهد که این الگوریتم ممکن است به همان اندازه برخی از الگوریتم‌های دیگر تعمیم‌پذیر نباشد.

الگوریتم جنگل تصادفی دقت آموزشی کامل (۱۰۰ درصد) و دقت آزمایشی بالا (۹۴/۷۹ درصد) دارد. جنگل تصادفی یک روش مجموعه‌ای است که معمولاً توانایی خوبی در تعمیم‌پذیری دارد. دقت آموزشی کامل با توجه به طبیعت الگوریتم غیرمنتظره نیست و دقت آزمایشی بالا نشان می‌دهد که الگوریتم مقاوم و تعمیم‌پذیر است.



دارد. الگوریتم KNN و SVM Support Vector عملکردهای متوسطی ارائه داده‌اند، با اما با AUC حدوداً ۰/۹۵ و تعداد خوبی از TP و TN، نتایج قابل قبولی را ارائه می‌دهند. به طور خلاصه، الگوریتم‌های LR و RF بهترین عملکرد را از خود نشان داده‌اند، در حالی که NB و DT نیازمند بهبود هستند با توجه به شکل ۷، در آنالیز صحت الگوریتم‌های مختلف یادگیری ماشین بر روی داده‌های دو کلاس «با تومور» و «بدون تومور»، مشاهده می‌شود که تعدادی از الگوریتم‌ها اختلافات و شباهت‌هایی در صحت تشخیص دارند. الگوریتم‌های رگرسیون لجستیک و ماشین بردار پشتیبان (SVM) به ترتیب صحت برای

کلاس «بدون تومور» ۹۷ و ۹۵ درصد، و برای کلاس «با تومور» به ترتیب ۹۹ و ۹۶ درصد دارند. همچنین، الگوریتم‌های نایب بیز (NB) و درخت تصمیم (DT) در مقایسه با سایر الگوریتم‌ها دارای صحت برای کلاس «بدون تومور» به ترتیب ۶۷ و ۸۰ درصد، و برای کلاس «با تومور» به ترتیب ۷۴ و ۸۵ درصد هستند. از طرف دیگر، الگوریتم‌های K- نزدیکترین همسایه (KNN) و جنگل تصادفی (RF) صحت برای کلاس «بدون تومور» به ترتیب ۷۷ و ۹۵ درصد، و برای کلاس «با تومور» به ترتیب ۹۷ و ۹۵ درصد دارند. بنابراین، الگوریتم‌های SVM و RF به عنوان انتخاب‌های مناسب برای تشخیص هر دو کلاس معرفی می‌شوند.



شکل ۷: نمایش نمونه تصاویر MRI پس از اتمام آزمایش

در این میان، الگوریتم‌های NB، KNN و DT نیز با امتیاز F1 کمتری برای کلاس‌های «بدون تومور» و «با تومور» عمل می‌کنند. این امتیازها به ترتیب برابر با (۸۲، ۸۶ درصد)، (۶۸، ۷۲ درصد) و (۸۲، ۸۳ درصد) بوده و نشان می‌دهند که عملکرد این الگوریتم‌ها در تمایز بین دو کلاس مورد نظر کمتر از الگوریتم‌های LR، SVM و RF است. صحت و امتیاز F1 الگوریتم‌ها، ارتباط مستقیم با توانایی تشخیص و تمایز بین کلاس‌ها دارد، و این نتایج می‌تواند در انتخاب و استفاده از الگوریتم‌های مناسب برای مسائل خاص در حوزه یادگیری ماشین ارزشمند واقع شود. شش الگوریتم در دیدن ۳۰ تصویر آزمایشی مشارکت و مورد ارزیابی قرار گرفتند (شکل ۷). الگوریتم بیز ساده ۳ تصویر، الگوریتم K- نزدیکترین همسایه دو تصویر و الگوریتم درخت تصمیم یک تصویر را اشتباه پیش‌بینی نمودند. این مشارکت همزمانی نشان می‌دهد که شش الگوریتم در ۲۴ تصویر نظر درستی را دادند و اجماع این نظر مشارکی مهم است. انتخاب تعداد تصاویر آزمایشی کم، خطای الگوریتم‌ها، تعمیم‌پذیری، نیاز به داده‌های بیشتر و تأثیر پیچیدگی الگوریتم می‌تواند در پیش‌بینی نوع تصاویر توسط الگوریتم‌ها مؤثر باشد.

در جدول ۱ میانگین صحت، بازخوانی و امتیاز F1 و همچنین دقت آزمایشی و سطح زیر منحنی راک این تحقیق با مقاله (۱۱) مورد مقایسه قرار گرفت و الگوریتم رگرسیون لجستیک، جنگل تصادفی و ماشین بردار پشتیبان نتایج بالاتری را از خود نشان می‌دهند. در مقاله (۲۴) که برای تشخیص سرطان مزوتلیوما بدخیم از الگوریتم‌های SVM، LR، KNN، NB، RF، DT بصورت باینری استفاده نمود، مقادیر صحت، بازخوانی، امتیاز F1 و دقت از ۷۰ تا ۱۰۰ درصد تعیین شد که کمترین مقدار الگوریتم NB و بیشترین مقدار را الگوریتم RF به خود اختصاص داد و مقادیر نتایج این تحقیق از ۷۰ تا ۹۸ درصد می‌باشند که کمترین مقدار را NB و بیشترین مقدار را LR به خود اختصاص داد (مقدار اختلاف ۲ درصد می‌باشد).

بر اساس (۲۳) بازخوانی یک معیار مهم در ارزیابی عملکرد الگوریتم‌های یادگیری ماشین است که نشان‌دهنده توانایی الگوریتم در شناسایی تمام نمونه‌های مثبت واقعی است. در اینجا، نتایج بازخوانی برای شش الگوریتم یادگیری ماشین برای تشخیص دو کلاس «بدون تومور» و «با تومور» ارزیابی شده است.

در این مقایسه، بازخوانی الگوریتم LR برای کلاس «بدون تومور» ۹۹ درصد و برای کلاس «با تومور» ۹۷ درصد است. در حالی که الگوریتم SVM دارای بازخوانی برابر با ۹۶ درصد برای کلاس «بدون تومور» و ۹۵ درصد برای کلاس «با تومور» می‌باشد. الگوریتم NB با مقادیر بازخوانی ۷۸ و ۶۲ درصد برای دو کلاس و الگوریتم KNN با مقادیر ۹۷ درصد برای کلاس «بدون تومور» و ۷۱ درصد برای کلاس «با تومور» ارزیابی شده‌اند. همچنین، الگوریتم RF دارای بازخوانی برابر با ۹۵ درصد برای هر دو کلاس است و الگوریتم DT دارای بازخوانی ۸۶ درصد برای کلاس «بدون تومور» و ۷۹ درصد برای کلاس «با تومور» است.

بر اساس نتایج الگوریتم‌های LR و SVM دارای بازخوانی بالاتری برای شناسایی تمام نمونه‌های مثبت واقعی هستند. با این حال، الگوریتم NB برای کلاس «با تومور» دارای بازخوانی پایین‌تری است که ممکن است به دلیل نیاز به بهبود عملکرد و توانایی تشخیص درست باشد. از این جهت، بازخوانی معیار مهمی است که تأثیر مستقیمی بر عملکرد الگوریتم در تشخیص دو کلاس مورد بررسی دارد و می‌تواند معیار مناسبی برای انتخاب الگوریتم مناسب در مسائل تشخیصی و پیش‌بینی باشد.

با بررسی اطلاعات، مشخص شد که الگوریتم‌های SVM، LR و RF دارای امتیاز F1 برابری به ترتیب (۹۸، ۹۸ درصد)، (۹۵، ۹۵ درصد) و (۹۵، ۹۵ درصد) برای دو کلاس «بدون تومور» و «با تومور» هستند، این نتایج نشان می‌دهند که عملکرد این سه الگوریتم در تشخیص دو کلاس مورد مقایسه یکسان بوده و امکان تمایز بهتری بین آن‌ها وجود ندارد.

جدول ۱: مقایسه معیارهای عملکرد الگوریتم‌ها با مرجع ۱۱

الگوریتم	دقت		صحت		بازخوانی		امتیاز F1		سطح زیر منحنی راک	
	این کار	مرجع ۱۱	این کار	مرجع ۱۱	این کار	مرجع ۱۱	این کار	مرجع ۱۱	این کار	مرجع ۱۱
SVM	۰/۹۵۴	۰/۸۵	۰/۹۵۵	۰/۷۹۸	۰/۹۵۵	۰/۷۵۴	۰/۹۵	۰/۷۷۵	۰/۹۵	۰/۸۸۸
LR	۰/۹۷۷	۰/۸۴۶	۰/۹۸	۰/۸	۰/۹۸	۰/۷۳۰	۰/۹۸	۰/۷۶۳	۰/۹۸	۰/۸۸۶
KNN	۰/۸۴۱	۰/۸۴۵	۰/۸۷	۰/۷۵۶	۰/۸۴	۰/۸۰۴	۰/۸۴	۰/۷۷۹	۰/۸۴	۰/۹۲
NB	۰/۷۰	۰/۷۶۹	۰/۷۰۵	۰/۶۷۵	۰/۷۰	۰/۶۲۷	۰/۷۰	۰/۶۵۰	۰/۷۰	۰/۷۵۵
DT	۰/۸۲۵	۰/۸۶۵	۰/۸۲۵	۰/۷۸۸	۰/۸۲۵	۰/۸۲۵	۰/۸۲۵	۰/۸۴۷	۰/۸۲	۰/۸۴۷
RF	۰/۹۴۷	۰/۹۰۸	۰/۸۵	۰/۸۱۱	۰/۹۵	۰/۹۵۲	۰/۹۵	۰/۹۷۵	۰/۹۵	۰/۹۷۵

الگوریتم یادگیری ماشین بررسی شد و امکان وجود الگوریتم‌های دیگری با عملکرد دیگری وجود دارد. بنابراین، برای بهبود دقت تشخیص و توسعه‌ی مدل‌های بهتر، لازم است که تحقیقات بیشتری در این زمینه انجام شود.

### تشکر و قدردانی

این مقاله منتج از طرح تحقیقاتی رشته‌ی فیزیک پزشکی با شماره ۴۰۳۰۰۰۰۱۰ می‌باشد که در دانشکده‌ی علوم پزشکی شوشتر به تصویب رسیده و با حمایت مالی بدون هزینه به انجام رسیده است. بدین‌وسیله از زحمات شورای پژوهشی و کمیته‌ی اخلاق تقدیر و تشکر می‌شود.

### نتیجه‌گیری

هدف از طرح پژوهشی حاضر، استفاده از مشارکت چندین الگوریتم بطور همزمان در تشخیص نوع تصویر با تومور و یا بدون تومور می‌باشد. با توجه به بررسی‌های انجام شده، الگوریتم رگرسیون لجستیک و جنگل تصادفی بهترین عملکرد را از خود نشان دادند. الگوریتم‌های نایب بیز و درخت تصمیم نیازمند بهبود هستند. الگوریتم‌های رگرسیون لجستیک و ماشین بردار پشتیبان بازخوانی بالاتری داشتند. معیار F1 توازن خوبی بین صحت و بازخوانی به نمایش گذاشت. انتخاب مناسب تعداد تصاویر آزمایشی نقش مهمی در عملکرد الگوریتم‌ها داشت. در نهایت باید توجه داشت که مطالعه‌ی حاضر، تنها بر اساس مجموعه‌ای از تصاویر MRI انجام شده است و نتایج ممکن است بر اساس داده‌های مختلف تغییر کنند. همچنین باید در نظر داشت که در تحقیق تنها شش

### References

1. Krishnapriya S, Karuna Y. Pre-trained deep learning models for brain MRI image classification. *Front Hum Neurosci* 2023; 17: 1150120.
2. Asad R, Rehman Su, Imran A, Li J, Almuhaimeed A, Alzahrani A. Computer-aided early melanoma brain-tumor detection using deep-learning approach. *Biomedicine* 2023; 11(1): 184.
3. Reza AW, Hossain MS, Wardifil MA, Farzana M, Ahmad S, Alam F, et al. A CNN-Based strategy to classify MRI-based brain tumors using deep convolutional network. *Appl Sci* 2023; 13(1): 312.
4. Thamarai M, Dhivyaa S. Analysis of Brain Tumor Classification using Pre-Trained CNN models. *Heliyon* 2024; 10(17): e36773.
5. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annual review of biomedical engineering. Annu Rev Biomed Eng* 2017; 19: 221-48.
6. Kuraparthi S, Reddy MK, Sujatha C, Valiveti H, Duggineni C, Kollati M, et al. Brain tumor classification of MRI images using deep convolutional neural network. *Traitement du signal* 2021; 38(4): 1171-9.
7. Nayeem MAH, Shakil MH, Afrin S, Shanto SA, Mumu SJ, Shanto MMHJJJoR, et al. A Deep Learning Based Classification Model for the Detection of Brain Tumor using MRI. *Journal of Research and Innovation in Applied Science* 2022; 7(9): 37-42.
8. Sharma K, Kaur A, Gujral S. Brain Tumor Detection based on Machine Learning Algorithms. *International Journal of Computer Applications* 2014; 103(1): 7-11.
9. Remzan N, Karim T, Farchi A. Brain tumor classification in magnetic resonance imaging images using convolutional neural network. *International Journal of Electrical and Computer Engineering*. 2022; 12(6): 6664-74.
10. Murali E, Meena K. A novel approach for classification of brain tumor using R-CNN. *International Journal of Engineering Applied Sciences and Technology* 2019; 4(4): 360-4.
11. Ghosh A, Kole A. A comparative study of enhanced machine learning algorithms for brain tumor detection and classification. [online] 2021. Available from: <https://www.techrxiv.org/doi/full/10.36227/techrxiv.16863136.v1>
12. Akinyelu AA, Zaccagna F, Grist JT, Castelli M, Rundo LJJo. Brain tumor diagnosis using machine learning, convolutional neural networks, capsule neural networks and vision transformers, applied to MRI: a survey. *J Imaging* 2022; 8(8): 205.
13. Praveena S, Singh SP, Kumar BS. Machine learning mechanism for segmentation, progressive assessment and prediction of brain tumor growth. *International journal of health sciences. Health Sciences* 2022; 6(S2): 7696-709.
14. Srinivas B, Rao SG. A hybrid CNN-KNN model for MRI brain tumor classification. *International Journal of Recent Technology and Engineering* 2019; 8(2): 5230-5.
15. Kaur P, Raja M, Lone A, Kaur M, Sansoya N. Exploring machine learning approaches for predicting brain tumors: a comparative study. *International Journal of Membrane Science and Technology* 2023; 10(5): 364-74.
16. Wang Z, Xiao X, He K, Wu D, Pang P, Wu T. A study of MRI-based machine-learning methods for glioma grading. *Int J Radiat Res* 2022; 20(1): 115-20.
17. Seethalakshmi BJB. Brain Tumor Malignancy Prediction Using Machine Learning Techniques. *Irish Interdisciplinary Journal of Science & Research* 2024; 8(2): 86-93.
18. Güler M, Namlı E. Brain Tumor Detection with Deep Learning Methods' Classifier Optimization Using Medical Images. *Appl Sci* 2024; 14(2): 642.
19. Wasule V, Sonar P, editors. Classification of brain MRI using SVM and KNN classifier. *Proceedings of the 3rd International Conference on Sensing, Signal Processing and Security (ICSSS); 2017 4-5 May; 2017.*

20. Yu Z, He Q, Yang J, Luo M. A Supervised ML Applied Classification Model for Brain Tumors MRI. *Front Pharmacol* 2022; 13: 884495.
21. Lamrani D, Cherradi B, el Gannour O, Bouqentar M, Bahatti L. Brain tumor detection using MRI Images and convolutional neural network. *International Journal of Advanced Computer Science and Applications* 2022; 13(7).
22. Imam R, Alam MT. Optimizing Brain Tumor Classification: A Comprehensive Study on Transfer Learning and Imbalance Handling in Deep Learning Models. In: Cuzzolin F, Sultana M (Editors). *Epistemic Uncertainty in Artificial Intelligence*. Pittsburgh, PA: 2023.
23. Chen H, Wang N, Du X, Mei K, Zhou Y, Cai G. Classification Prediction of Breast Cancer Based on Machine Learning. *Comput Intell Neurosci* 2023; 2023: 6530719.
24. Win KY, Maneerat N, Choomchuay S, Sreng SKH. Suitable Supervised Machine Learning Techniques For Malignant Mesothelioma Diagnosis. *Proceedings of the 11<sup>th</sup> Biomedical Engineering International Conference (BMEiCON)*; 2018 21-24 Nov; 2018.

## Evaluation of Machine Learning Algorithms for Predicting Tumor and Non-tumor Brain Mri Images

Barat Barati<sup>1</sup>, Maryam Erfaninejad<sup>2</sup>, Sima Hashemi<sup>3</sup>, Nahid Chegeni<sup>4</sup>, Mohsen Arshadi<sup>5</sup>

### Original Article

#### Abstract

**Background:** Early diagnosis of brain tumors using MRI and artificial intelligence algorithms is fundamental in improving treatment results. MRI images serve as the primary tool for identifying brain tumors. This study aims to evaluate machine learning algorithms for diagnosing brain tumors and non-tumors using MRI images.

**Methods:** From kaggle.com a total of 2400 MRI images were collected, and a pre-processing step was performed on them. Algorithms such as logistic regression, decision tree, random forest, simple Bayes method, support vector machine, and K nearest neighbor were also implemented on the images.

**Findings:** After applying all the algorithms, the values of training accuracy, test accuracy, accuracy, readability, F1 score, confusion matrix, and the area under the rocking curve were obtained to evaluate the performance criteria.

**Conclusion:** The investigations indicated that logistic regression and random forest algorithms performed the best. Naive Bayes and decision tree algorithms need improvement.

**Keywords:** MRI images; Diagnosis; Machine learning; Logistic models; Random forest

**Citation:** Barati B, Erfaninejad M, Hashemi S, Chegeni N, Arshadi M. **Evaluation of Machine Learning Algorithms for Predicting Tumor and Non-tumor Brain MRI Images.** J Isfahan Med Sch 2024; 42(778): 674-86.

1- Assistant Professor, Department of Radiology Technology, School of Medical Sciences, Shoushtar Faculty of Medical Sciences, Shoushtar, Iran

2- Assistant Professor, Department of Basic Sciences, School of Medical Sciences, Shoushtar Faculty of Medical Sciences, Shoushtar, Iran

3- Assistant Professor, Department of Nursing Education, School of Medical Sciences, Shoushtar Faculty of Medical Sciences, Shoushtar, Iran

4- Associate Professor, Department of Medical Physics, School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

5- Shoushtar Faculty of Medical Sciences, Shoushtar, Iran

**Corresponding Author:** Barat Barati, Assistant Professor, Department of Medical Physics, School of Medical Sciences, Shoushtar Faculty of Medical Sciences, Shoushtar, Iran; Email: Barati943@yahoo.com