سامانه‌های پردازشی و ارتباطی چندرسانه‌ای هوشمند
Intelligent Multimedia Processing and Communication Systems (IMPCS)

Spring 2024, 5 (1), 37-45
DOR:

Received: 6 Jan 2024
Accepted: 14 Feb 2024

مقاله پژوهشی

An Online Group Feature Selection Algorithm using Mutual Information

Maryam Rahmaninia[1*], Sondos Bahadori[2]

1. Assistant Professor, Department of Computer Engineering, Qasreshirin Branch, Islamic Azad University, Qasreshirin, Iran. * *Corresponding Author*, ma.rahmaninia@gmail.com
2. Assistant Professor, Department of Computer Engineering, Ilam Branch, Islamic Azad University, Ilam, Iran. sondos.bahadori@iau.ac.ir

Abstract

Introduction: In the area of big data, the dimension of data in many fields are increasing dramatically. To deal with the high dimensions of training data, online feature selection algorithms are considered as very important issue in data mining. Recently, online feature selection methods have attracted a lot of attention from researchers. These algorithms deal with the process of selecting important and efficient features and removing redundant features without any pre-knowledge of the set of features. Despite all the progress in this field, there are still many challenges related to these algorithms. Among these challenges, we can mention scalability, minimum size of selected features, sufficient accuracy and execution time. On the other hand, in many real-world applications, features are entered into the dataset in groups and sequentially. Although many online feature selection algorithms have been presented so far, but none of them have been able to find trade of between these criteria.

Method: In this paper, we propose a group online feature selection method with feature stream using two new measures of redundancy and relevancy using mutual information theory. Mutual information can compute linear and non-linear dependency between the variables. With the proposed method, we try to create a better tradeoff between all the challenges.

Results: In order to show the effectiveness of the proposed online group feature selection method, a number of experiments have been conducted on six large multi-label training data sets named ALLAML, colon, SMK-CAN-187, credit-g, sonar and breast-cancer in different applications and 3 online group feature selection algorithms named FNE_OGSFS ،Group-SAOLA and OGSFS which are presented recently. Also, 3 evaluation criteria including average accuracy using KNN (k - nearest neighborhood) , SVM (Support Vector Machine) and NB (Naïve Bayesian) classifiers, number of selected features and executing time were used as criteria for comparing the proposed method. According to the obtained results, the proposed algorithm has obtained better results in almost of cases compared to other algorithms which it shows the efficiency of the proposed method.

Discussion: In this paper, we will show that proposed online group feature selection method will achieve better performance by considering label group dependency between the new arrival features.

Keywords: Online Group Feature selection algorithm, feature stream datasets, mutual information.

سامانه‌های پردازشی و ارتباطی چندرسانه‌ای هوشمند
Intelligent Multimedia Processing and Communication Systems (IMPCS)

یک الگوریتم انتخاب ویژگی برخط گروهی با استفاده از نظریه اطلاعات متقابل

تاریخ دریافت: ۱۶/ ۱۰/ ۱۴۰۲
تاریخ پذیرش: ۲۵/ ۱۱/ ۱۴۰۲

دورهٔ پنجم، بهار ۱۴۰۳
شمارهٔ اول، صص: ۳۷-۴۵

مریم رحمانی‌نیا[1*]، سندس بهادری[2]

۱. استادیار، گروه کامپیوتر، مرکز قصرشیرین، دانشگاه آزاد اسلامی، قصرشیرین، ایران. (نویسنده مسئول) ma.rahmaninia@gmail.com
۲. استادیار، گروه کامپیوتر، واحد ایلام، دانشگاه آزاد اسلامی، ایلام، ایران. Sondos.bahadori@iau.ac.ir

چکیده

امروزه ویژگی‌ها به‌صورت پیوسته و یکی‌یکی یا گروهی به مجموعه داده‌ها اضافه‌می‌شوند. این امر باعث افزایش بی‌رویه ابعاد داده‌های آموزشی می‌شود. روش‌های انتخاب ویژگی برخط، با انتخاب یک زیر مجموعه مهم از میان مجموعه ویژگی‌های دیده‌شده تا آن لحظه اندازه داده‌های آموزشی را کاهشمی‌دهند و در نتیجه عملکرد الگوریتم‌های یادگیری ماشین را بهبود می‌بخشند. اما به‌دلیل عدم‌قطعیت و ابهام در داده‌های آموزشی، روش‌های انتخاب ویژگی برخط موجود معمولاً از مشکلاتی مانند مقیاس‌پذیری، دقت پایین و زمان اجرای بالا رنج‌می‌برند. همچنین بسیاری از این روش‌ها توانایی انتخاب ویژگی‌های مهم هنگامی که ویژگی‌ها به‌صورت گروهی به مجموعه داده‌ها اضافه‌می‌شوند را ندارند. بنابراین در این مقاله، یک روش انتخاب ویژگی برخط گروهی بر اساس نظریه اطلاعات متقابل ارائه‌داده‌ایم. در روش پیشنهادی در ابتدا مقدار ارتباط هر یک از ویژگی‌های موجود در گروه تازه‌وارد شناسایی و سپس مقدار افزونگی ویژگی‌های انتخابی در گروه تازه‌وارد و مجموعه ویژگی‌هایی که قبلاً انتخاب‌شده محاسبه و ویژگی‌های اضافی حذف‌می‌شوند. نتایج به‌دست‌آمده توسط الگوریتم پیشنهادی در مقایسه با ۳ الگوریتم انتخاب ویژگی برخط گروهی دیگر ارزیابی‌شده‌است. طبق نتایج حاصل، الگوریتم پیشنهادی کارایی بالاتری نسبت به سایر الگوریتم‌ها به‌دست‌آورده‌است.

واژه‌های کلیدی: الگوریتم‌های انتخاب ویژگی گروهی برخط، داده‌های با جریان برخط، اطلاعات متقابل.

۱. مقدمه

امروزه با حجم وسیعی از اطلاعات سروکار داریم. این اطلاعات از منابع مختلفی مانند آزمایشگاه، تصاویر ماهواره‌ای، بیولوژیک و بینایی ماشین به‌دست آمده‌است[۱-۳]. از طرف دیگر، اندازه و حجم بالای این داده‌ها کارایی بسیاری از الگوریتم‌های یادگیری ماشین را تحت شعاع قرارداده و باعث‌شده بسیاری از این روش‌ها عملکرد خود را از دست بدهند[۴, ۵]. روش‌های انتخاب ویژگی، یک رویکرد مهم به‌منظور کاهش تعداد ویژگی‌ها و در نتیجه کاهش اندازه داده‌های آموزشی است [۶]. در این روش‌ها معنای اولیه ویژگی‌ها بدون تغییر مانده و فقط تعدادی از ویژگی-های مهم انتخاب‌شده و ویژگی‌های اضافی و نامرتبط از مجموعه ویژگی‌ها حذف‌می‌شود. تاکنون روش‌های انتخاب ویژگی مختلفی ارائه‌شده‌است که می‌توان آن‌ها را در سه دسته فیلتر، پروانه‌ای و جاسازی‌شده دسته-بندی‌کرد. در روش‌های فیلتر با استفاده از معیارهای مختلفی چون بهره اطلاعاتی، نرخ فیشر[۷]، همبستگی[8] ، فاصله اقلیدسی [۹]، اطلاعات متقابل [۱۰] و غیره مقدار کارایی ویژگی‌ها اندازه‌گیری‌شده و بعد از مرتب‌سازی آن‌ها، k تا از بهترین ویژگی‌ها که در ابتدای صف به‌وجودآمده قرارگرفته‌اند انتخاب‌می‌شوند. در مقابل، روش‌های پوششی با استفاده از سیاست‌های مختلف زیر مجموعه‌های مختلف از ویژگی‌ها را انتخاب‌می-کنند و بر اساس معیارهای مختلف آن‌ها را ارزیابی و رتبه‌دهی می‌کند[۷, ۱۱]. سپس هرکدام از زیر مجموعه ها که دارای مطلوبیت بالاتری بود انتخاب‌می‌شوند. در نهایت روش‌های انتخاب ویژگی از نوع جاسازی‌شده، با ادغام دو رویکرد فیلتر و پوششی به‌طور همزمان از مزایای هر دو روش استفاده‌کرده‌است[۱۲, ۱۳] .

اگرچه تاکنون روش‌های انتخاب ویژگی معرفی‌شده در هر کدام از این سه دسته توانسته‌است مسائل زیادی را برطرف‌کند اما هنوز چالش-های زیادی وجوددارد که مرتفع نشده‌است. یکی از این چالش‌ها کامل-شدن تدریجی داده‌های آموزشی است. امروزه به‌دلیل رشد تکنولوژی داده‌های آموزشی به‌تدریج و با گذر زمان اضافه‌می‌شوند. به‌عنوان مثال شبکه‌های اجتماعی چون فیسبوک و توییتر روزانه با هزاران کلمه جدید روبرو و مخزن داده‌های مربوط به این شبکه‌ها کامل‌تر می‌شوند[۱۴]. از طرف دیگر با اضافه‌شدن هرکدام از این کلمات، اضافی بودن هر یک از کلمات قبلی و یا حتی کلمه‌های جدید ممکن است رخ‌دهد. گاهی اوقات نیز، ممکن است کلمات جدید هیچ ربطی به هیچ‌یک از دامنه‌های کاربردی نداشته‌باشد. به همین دلیل، در اینگونه داده‌های آموزشی که ویژگی‌ها به‌صورت تدریجی به داده‌ها اضافه‌می‌شوند نیز ممکن است تعداد ویژگی‌ها آنقدر زیاد شود که باعث افزایش بی‌رویه حجم داده‌های آموزشی شود. در این مواقع به‌کار بردن روش‌های انتخاب ویژگی برخط، می‌تواند باعث کاهش حجم داده‌ها و بهبود کیفیت الگوریتم‌های یادگیری ماشین شود[۱۵, ۱۶]. این الگوریتم‌ها ازآنجاکه در محیط‌های برخط استفاده‌می‌شوند بهتر است در حداقل‌ترین زمان ممکن کارایی یک ویژگی را از لحاظ افزونگی و یا ارتباط با برچسب کلاس بررسی‌کنند. همچنین در هر لحظه زمانی تعداد ویژگی‌های انتخابی باید حداقل باشد. از طرف دیگر چون هیچ‌گونه اطلاعاتی از داده‌هایی که در آینده قرار است به مجموعه آموزشی واردشوند در دسترس نیست، پس بهتر است بدون هیچ دانش پیش زمینه‌ای به فرایند انتخاب ویژگی‌ها بپردازد. تاکنون الگوریتم‌های انتخاب ویژگی برخط متعددی ارائه‌شده‌است که هر کدام سعی در برطرف کردن یک یا چند چالش دارد [20-16]. اگر چه همه این روش‌ها تا حدودزیادی توانسته‌است شرایط یک الگوریتم انتخاب ویژگی برخط با ویژگی جریانی را برآورده‌کند. اما در همه آن‌ها فرض بر این است که ویژگی‌ها به‌صورت منفرد و یکی‌یکی به مجموعه داده‌ها واردمی‌شوند. در دنیای واقعی گاهی اوقات ویژگی‌ها به‌صورت گروهی به مجموعه داده اضافه‌می‌شوند. به‌عنوان مثال، برای یک شخص، آیتم‌های مختلف آزمایش گروه‌های مختلفی از ویژگی‌ها مانند معاینه معمول خون، معاینه معمول ادرار و غیره را تولیدمی‌کند [۲۱]. اخیراً چندین روش انتخاب ویژگی برخط گروهی در جریان داده‌ها ارئه‌شده‌است [۲۲-۲۴]. اما به‌دلیل عدم‌قطعیت و وجود ابهام در داده‌های آموزشی هیچ‌کدام از این روش‌ها نتواسته‌اند تاکنون حدّواسط مطلوبی میان چالش‌هایی چون مقیاس‌پذیری، زمان اجرا، دقت و تعداد ویژگی‌های انتخاب‌شده، ایجادکنند.

دراین مقاله قصد داریم یک الگوریتم انتخاب ویژگی برخط در داده-های آموزشی جریانی که در آن ویژگی‌ها به‌صورت گروهی و برخط به مجموعه داده‌ها اضافه‌می‌شوند، ارائه‌دهیم که به‌منظور ارزیابی ویژگی‌های درون یک گروه و نیز، گروه‌های مختلف از ویژگی‌ها از معیار اطلاعات متقابل استفاده‌شده‌است. اطلاعات متقابل در نظریه اعداد یک معیار کارا و مؤثر در جهت کشف وابستگی‌های خطی و غیرخطی میان متغیرها به‌حساب‌می‌آید.

ساختار کلی مقاله پس از مقدمه‌ای که گذشت، شامل پیشینه تحقیق و ذکر کارهایی است که تاکنون در این زمینه انجام‌شده‌است (بخش دوم)، بخش سوم و چهارم به‌ترتیب شامل روش پیشنهادی و نتایج آزمایشگاهی به همراه تجزیه و تحلیل نتایج است. بخش پنجم نیز شامل نتیجه‌گیری و کارهای آینده است.

۲. پیشینه تحقیق

۱،۲. پیشینه تحقیق

به‌طور کلی می‌توان براساس نوع داده‌های آموزشی، الگوریتم‌های انتخاب ویژگی مختلف را در دو دسته ایستا و برخط دسته‌بندی‌کرد. الگوریتم-های انتخاب ویژگی ایستا فرض‌می‌کنند که در لحظه شروع الگوریتم همه داده‌ها در دسترس است [۲۵-۲۹]. در صورتی که در دنیای واقعی این-گونه نیست و ویژگی‌ها به‌صورت پی‌درپی و تدریجی به مجموعه داده‌ها اضافه‌می‌شوند[۱۴, ۲۱]. در چنین داده‌هایی انتظار کامل‌شدن داده‌های آموزشی قبل از انجام فرایند انتخاب ویژگی، عملاً بی‌فایده است. به همین دلیل، الگوریتم‌های انتخاب ویژگی برخط مطرح‌شد. در این الگوریتم‌ها هر لحظه که داده جدیدی به مجموعه داده‌ها اضافه‌می‌شود فرایند انتخاب ویژگی انجام‌می‌شود. بر این اساس که ویژگی‌ها به‌صورت فردی

یا گروهی به مجموعه داده‌ها اضافه‌می‌شوند می‌توان این الگوریتم‌ها را در دو دسته فردی و گروهی برخط دسته‌بندی‌کرد. در الگوریتم‌های انتخاب ویژگی فردی برخط فرض بر این است که ویژگی‌ها به‌صورت یکی‌یکی به مجموعه داده‌ها اضافه‌می‌شوند. در ادامه چند الگوریتم انتخاب ویژگی برخط فردی را معرفی‌می کنیم. برای اولین بار در [۲۰] نویسندگان یک روش جدید به منظور حل مسأله انتخاب ویژگی برخط فردی را بر اساس رویکرد کاهش گرادیان خطا به نام grafting پیشنهادکردند. این روش با هر دو مدل پیش‌بینی‌کننده خطی و غیرخطی به‌خوبی عمل‌می‌کند و هم برای طبقه‌بندی و هم برای رگرسیون می‌تواند استفاده‌شود. در [30] یک الگوریتم انتخاب ویژگی برخط فردی دیگر به نام Alpha-investing و نسخه تکامل‌یافته آن براساس رگرسیون جریانی برای مسأله انتخاب ویژگی برخط ارائه‌شد. این الگوریتم‌ها، می‌توانند تعداد نامشخص و حتی نامتناهی از ویژگی‌ها را سازماندهی‌کند. اما ازآنجاکه ویژگی‌ها را دوباره ارزیابی‌نمی‌کند افزونگی میان ویژگی‌ها را درنظرنمی‌گیرد. در [۱۵] یک چهارچوب انتخاب ویژگی با دو الگوریتم به نام OSFS و fast-OSFS ارائه‌شد که با استفاده از نظریه احتمالات شرطی به فرایند انتخاب ویژگی‌های مهم پرداخته‌است. این الگوریتم‌ها ویژگی‌ها را به سه دسته مرتبط قوی، مرتبط ضعیف و نامربوط با استفاده از نظریه احتمالات شرطی دسته‌بندی‌می‌کند. در [۳۱] دو الگوریتم انتخاب ویژگی برخط فردی به‌نام OFS و OFSp به‌ترتیب ورودی کامل و ورودی جزئی ارائه-شده‌است. در [۳۲] یک الگوریتم انتخاب ویژگی به نام SAOLA ارائه-شده‌است. این الگوریتم دو چالش اساسی ابعاد بالا و مقیاس‌پذیری را در الگوریتم‌های انتخاب ویژگی برخط بررسی‌کرده‌است که با استفاده از بررسی دو به دوی ویژگی‌ها بر اساس اطلاعات متقابل، به‌صورت برخط به فرایند انتخاب ویژگی می‌پردازد. در [۳۳] یک روش انتخاب ویژگی برخط دیگر نیز به نام OSFSMI برای فرایند انتخاب ویژگی‌های مهم و حذف ویژگی‌های افزونه در جریان داده‌ها با استفاده از نظریه اطلاعات متقابل شرطی ارائه‌شد.

در مقابل روش‌های انتخاب ویژگی گروهی برخط، به فرایند انتخاب ویژگی در حالتی که ویژگی‌ها به‌صورت گروهی به مجموعه داده‌ها اضافه‌می‌شود پرداخته‌می‌شود. در [۳۴] یک روش انتخاب ویژگی برخط گروهی به نام GFSSF ارائه شده‌ا ست که با ا ستفاده از نظریه اطلاعات متقابل و آنتروپی قادر به انتخاب ویژگی‌های مهم و کارآمد هم در سطح فردی و هم در سطح گروهی است. این الگوریتم دارای دو فاز درون گروهی و برون‌گروهی است و بر اساس ارتباط میان ویژگی‌ها، آن‌ها را انتخاب‌می‌کند. در [۲۳] یک روش انتخاب ویژگی برخط گروهی دیگر به نام OGFS ارائه‌شده‌است. این روش نیز شامل دو مرحله انتخاب درون‌گروهی برخط و انتخاب میان گروهی برخط است. در انتخاب برخط درون‌گروهی، ویژگی‌ها با ا ستفاده از معیار تجزیه و تحلیل طیفی و در انتخاب میان‌گروهی برخط، از مدل رگرسیون خطی به‌منظور انتخاب یک زیرمجموعه بهینه از ویژگی‌ها استفاده‌شده‌است. نویسندگان در [۲۴] یک روش دیگر به نام Group-SAOLA ارائه‌داده‌اند که نسخه گسترش‌یافته روش SAOLA است. در این روش در ابتدا براساس اطلاعات متقابل در فاز درون‌گروهی ویژگی‌های مرتبط و افزونه در گروه تازه‌وارد شنا سایی‌می شود. سپس در فاز برون‌گروهی میزان افزونگی و ارتباط گروه تازه‌وارد در مقایسه با سایر گروه‌ها محاسبه و گروه‌های اضافی شناسایی و از مجموعه حذف‌می‌شوند. در [۲۲]، یک روش انتخاب ویژگی به‌نام OGSFS ارائه شده‌ا ست که ویژگی‌ها را طی دو فاز درون گروهی و میان گروهی به‌طور همزمان با یکدیگر انتخاب‌می‌کند. در فاز درون گروهی ویژگی هایی از یک گروه که با هم تعامل‌دارند، شنا سایی شده و سپس در فاز انتخاب میان گروه، با ا ستفاده از روش تنظیم‌سازی و انتخاب متغیرهای الاستیک نت گروه‌های تأثیرگذارتر شنا سایی‌می شوند. در [۳۵] یک روش جدید دیگر به منظور انتخاب ویژگی جریان گروهی برخط به نام FNE-OGSFS با استفاده از آنتروپی همسایگی فازی و مجموعه‌های خشن همسایگی ارائه‌شده‌است که ویژگی‌های مهم در گروه‌ها را شنا سایی و انتخاب‌می‌کند. سپس، تمام ویژگی‌های انتخابی توسط مدل Lasso مورد ارزیابی مجدد قرارمی‌گیرند تا ویژگی‌های اضافی حذف‌شوند.

۲,۲ پیش زمینه

در نظریه احتمالات، اطلاعات متقابل بین دو متغیر تصادفی معیاری برای نشان‌دادن میزان وابستگی متقابل آن دو متغیر می‌باشد. مفهوم اطلاعات متقابل ذاتاً مرتبط با آنتروپی یک متغیر تصادفی که میزان اطلاعات موجود در یک متغیر تصادفی را نشان‌می‌دهد، است. آنتروپی یکی از معیارهای اندازه‌گیری میزان عدم‌قطعیت یک متغیر به‌حساب می‌آید که به‌ازای متغیر $X=[x_1,x_2,\dots x_n]$ به‌صورت زیر تعریف‌می-شود:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log p(x_i) \qquad (۱)$$

به‌طوری که $p(x_i)$ احتمال رخداد مقدار x_i از مجموعه مقادیر ممکن برای متغیر X را نشان‌می‌دهد. حال آنتروپی توأم دو متغیر تصادفی $X=[x_1,x_2,\dots x_m]$ و $Y=[y_1,y_2,\dots y_m]$ که لزوماً از هم مستقل نیستند، عبارت است از

$$H(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i,y_j)\log p(x_i,y_j) \qquad (۲)$$

$p(x,y)$ به احتمال توأم دو مقدار x و y اشاره‌می‌کند.

در نظریه اطلاعات، اطلاعات متقابل بین دو متغیر به مقداری از اطلاعات که بین دو متغیر به ا شتراک گذا شته شده‌ا ست، ا شاره‌می‌کند که آن را می‌توان با استفاده از رابطه زیر به‌دست‌آورد:

$$I(X,Y)=\sum_{x\in X}\sum_{y\in Y}p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \qquad (۳)$$

مقدار رابطه (۳) معمولاً مثبت اســت. اگر اطلاعات متقابل میان دو ویژگی زیاد باشــد به این معنی اســت که وابســتگی دو متغیر نیز زیاد است.

۳. راهکار پیشنهادی

در این بخش چهار چوب اصلی روش پیشنهادی با استفاده از معیار اطلاعات متقابل که در فصل قبلی معرفی‌شده ارائه‌می‌شود. بطور کلی می‌توان‌گفت الگوریتم پیشنهادی دارای دو فاز اصلی است. فاز اول به ازای ورود هر گروه G از ویژگی‌های تازه‌وارد اجرامی‌شود. در این فاز ابتدا ویژگی‌های مهم و مرتبط با برچسب کلاس شناسایی و سپس ویژگی-هایی که با برچسب کلاس مرتبط نیستند از گروه حذف‌می‌شوند. درصورتی‌که بعد از فاز اول، همچنان ویژگی در گروه G باقی‌مانده‌باشد، الگوریتم وارد فاز دوم می‌شود. در فاز دوم ویژگی‌های انتخابی از گروه G به مجموعه ویژگی‌های انتخابی از مراحل قبلی الگوریتم (مجموعه S') اضافه و سپس میزان افزونگی آن‌ها نسبت به یکدیگر و برچسب کلاس محاسبه و ویژگی‌های اضافی از مجموعه S' حذف‌می‌شوند. بعد از این فاز الگوریتم همچنان منتظر ورود گروه‌های جدید از ویژگی‌ها می‌ماند. ویژگی‌هایی که از فاز اول و فاز دوم حذف‌می‌شوند، در مراحل بعدی اجرای الگوریتم حساب‌نمی‌شوند. در ادامه دو معیار برای محاسبه میزان ارتباط و افزونگی ویژگی‌ها معرفی‌شده و سپس با استفاده از این معیارها مراحل فازهای الگوریتم پیشنهادی را با جزئیات بیشتر معرفی‌می‌کنیم. علاوه بر این شبه کد الگوریتم پیشنهادی در شکل ۱ آمده‌است.

تعریف ۱ (ارتباط ویژگی): میزان ارتباط و تأثیر یک ویژگی $F=[f_1,f_2,\ldots,f_n]$ با برچسب کلاس C با استفاده از رابطه زیر به-دست‌می‌آید:

$$I(F;C)=\sum_{f_i\in F}\sum_{c\in C}p(f_i,c)\log\frac{p(f_i,c)}{p(f_i)p(c)} \qquad (۴)$$

به‌طوری که $I(X,C)$ به اطلاعات متقابل میان ویژگی X و برچسب C که از رابطه ۲ می‌توان آن را محاسبه‌کرد، اشاره‌می‌کند.

تعریف ۲ (افزونگی ویژگی): ویژگی F_2 نسبت به ویژگی F_1 اضافی درنظرگرفته‌می‌شود اگر و تنها اگر

$$I(F_1;F_2)>I(F_1;C)\ ,\ I(F_2;C)<I(F_1;C) \qquad (۵)$$

فاز اول (درون گروهی): این فاز به ازای ورود یک گروه تازه وارد از ویژگی‌ها به نام G اجرامی‌شود. در قدم اول در این فاز، مقدار ارتباط هر کدام از ویژگی‌ها با برچسب کلاس با استفاده از تعریف ۱ محاسبه شده و آن ویژگی‌هایی که مقدار ارتباط آن‌ها با برچسب کلاس کمتر از یک مقدار آستانه است، از مجموعه حذف‌می‌شود و سپس سایر ویژگی‌ها به-صورت نزولی و براساس مقدار ارتباط اشان در لیست S به‌صورت نزولی مرتب‌می‌شوند (خطوط ۱ تا ۶ الگوریتم پیشنهادی در شکل ۱). حال بعد از این مرحله اگر در گروه G ویژگی باقی‌ماند الگوریتم وارد فاز دوم می‌شود. در غیراین‌صورت الگوریتم بدون اضافه‌کردن هیچ ویژگی جدیدی به مجموعه ویژگی‌های نهایی خاتمه‌می‌یابد.

فاز دوم (برون‌گروهی): این فاز هنگامی اجرامی‌شود که حداقل یک ویژگی در گروه تازه از راه رسیده G باقی‌مانده‌باشد. در این فاز در ابتدا ویژگی‌های انتخابی از فاز اول به‌صورت مرتب و نزولی بر اساس میزان ارتباطشان با برچسب کلاس به مجموعه ویژگی‌های انتخابی از مراحل قبلی الگوریتم(S') اضافه‌می‌شوند. سپس براساس تعریف ۲، میزان افزونگی ویژگی‌ها نسبت به یکدیگر محاسبه‌می‌شود (خطوط ۱۱ تا ۱۹ الگوریتم پیشنهادی در شکل ۱). به این‌صورت که همانند خط ۹ از ویژگی‌ای که سر لیست S' قراردارد (مثلاً Fi) شروع‌می‌کنیم و تمام ویژگی‌هایی که بعد از آن در لیست قراردارند را از لحاظ میزان افزونگی طبق تعریف ۲ ارزیابی‌می‌کنیم. اگر Fj بعد از آن در لیست S' قرار داشته-باشد و اطلاعات متقابل آن‌ها بیشتر از مقدار اطلاعات متقابل ویژگی Fi با برچسب کلاس باشد پس ویژگی Fj از مجموعه S' حذف‌می‌شود. این کار تا انتهای لیست S' انجام‌می‌شود. بعد از این‌که یک‌بار لیست S' به‌ازای ویژگی Fi پیمایش‌شد و تعدادی از ویژگی‌ها حذف‌شدند، مجدد دومین ویژگی که حال بعد از F_i در سر لیست S' قرار دارد انتخاب‌شده و ویژگی‌هایی که طبق تعریف ۲ با آن دارای افزونگی هستند شناسایی و از لیست S' حذف‌می‌شود. این عملیات تا جایی ادامه‌می‌یابد که برای تمام ویژگی‌ها یک‌بار لیست S' پیمایش‌شده باشد. در نهایت ویژگی‌های باقی‌مانده در S' به‌عنوان مجموعه ویژگی‌های نهایی انتخاب‌می‌شوند و الگوریتم منتظر ورود ویژگی‌های دیگر می‌ماند.

```
Input: new arrival group of features G
Output: the final selected feature set S

# Phase 1:
For each feature F in new arrival group of features G
    Calculate MI=I(F , C)
    If ( MI > threshold)
     Insert F into selected feature set S in descending order
    End if
End For
#Phase 2:
If (S is not empty)
insert the features in S into S' in descending order
Fj= First _ Element (S) // select the first feature in sorted list S
    Do
     Fi= Next _ Element (S , Fj) //find feature fi in S which is next to Fj
      Do
        If I( Fi , Fj)>I( Fi , C)
         Remove fi from S
         Fi= Next _ Element ( S , Fj)
        End If
      End Until (Fi==Null);
      Fj= Next _ Element (S, Fj)
    End Until (Fj==Null)
  End
 End
Return S
```

شکل ۱: شبه کد الگوریتم پیشنهادی

۴. نتایج آزمایشگاهی

در این بخش به بررسی کارایی روش پیشنهادی در مقایسه با سه الگوریتم انتخاب ویژگی برخط FNE_OGSFS [۳۵]، Group-SAOLA[۲۴] و OGSFS[۲۲] که در حالت گروهی به فرایند انتخاب ویژگی‌ها می‌پردازند و ۶ داده آموزشی معروف و پرکاربرد که مشخصات آن‌ها در جدول ۱ آمده، می‌پردازیم. برای این کار از سه معیار ارزیابی تعداد ویژگی‌های انتخابی هر الگوریتم، دقت سه دسته‌بندها KNN[36]، SVM [۳۷] و NB [۳۸] و زمان اجرا برای ارزیابی الگوریتم‌ها استفاده‌شده‌است. هرچقدر تعداد ویژگی‌های انتخاب‌شده و زمان اجرا کمتر باشد الگوریتم مطلوب‌تر است. از طرف دیگر هرچقدر دقت به‌دست‌آمده به‌ازای انواع دسته‌بندها برای یک الگوریتم بیشتر باشد پس مطلوبیت الگوریتم بیشتر است.

جدول ۱: داده‌های آموزشی

	تعداد ویژگی ها	تعداد نمونه های آموزشی	تعداد کلاس ها	کاربرد
ALLAML	7129	72	2	Bio
Colon	2000	62	2	Bio
SMK-CAN-18	19993	187	2	Microarray, Bio
credit-g	1000	20	2	credit risk
Sonar	208	60	2	signal
breast.cancer	286	9	2	Bio

برای راحتی کار، اندازه گروه‌ها در همه الگوریتم‌ها ۵ درنظرگرفته‌شده‌است. به این معنی که هر بار یک گروه ۵ تایی از ویژگی‌ها به مجموعه اضافه‌می‌شود. همچنین ازآنجاکه ویژگی‌ها با ترتیب‌های مختلفی به مجموعه اضافه‌می‌شوند پس، نتایج حاصل روی ۵ بار اجرای هر کدام از الگوریتم‌ها و به‌ازای ترتیب‌های مختلف اجراشده‌است. علاوه براین تمام آزمایش‌ها روی یک سیستم با سیستم عامل ویندوز ۱۱، RAM 8 و گرافیک RTX 3050 انجام‌شده‌است.

۱،۴. ارزیابی روش پیشنهادی از لحاظ دقت دسته‌بندی

در این بخش به ارزیابی روش پیشنهادی در مقایسه با سایر روش‌های انتخاب ویژگی [۳۵] FNE_OGSFS، Group-SAOLA [۲۴] و OGSFS [۲۲] از لحاظ دقت دسته‌بندی می‌پردازیم. در ادامه نتایج به‌دست‌آمده برای هرکدام از این دسته‌بندها و داده‌های آموزشی به همراه الگوریتم‌های انتخاب ویژگی برخط گروهی در جدول‌های ۲ تا ۴ نشان‌داده‌شده‌است. در هر جدول سطرها نشان‌دهنده یک داده آموزشی و ستون‌ها مربوط به یک الگوریتم انتخاب ویژگی برخط گروهی است. همچنین بیشترین دقت به‌دست‌آمده به‌ازای هر داده آموزشی و هر الگوریتم انتخاب ویژگی به‌صورت ضخیم و زیر خط‌دار نشان‌داده‌شده‌است. نتایج به‌دست‌آمده به‌ازای دسته‌بند SVM در جدول شماره ۲ نشان‌داده‌شده‌است. همان‌طور که نتایج نشان‌می‌دهد ویژگی‌های انتخاب‌شده به‌ازای الگوریتم پیشنهادی و همه داده‌های آموزشی همواره به‌جز یک مورد بیشترین دقت را به‌دست آورده‌است. به‌عنوان مثال، به‌ازای داده آموزشی ALLAML، دقت به‌دست‌آمده به‌ازای الگوریتم پیشنهادی،FNE-OGSFS، Group-SAOLA و OGSFS به‌ترتیب ۹۸,۹۸، ۲۱,۷۱، ۶۲,۵۸ و ۷۰,۷۳ است. که نشان‌می‌دهد الگوریتم پیشنهادی تقریباً ۷۸٪، ۳۶٪ و ۲۸٪ نسبت به سایر الگوریتم‌ها بهتر عمل‌کرده‌است. در سطر آخر جدول ۲، مقادیر میانگین به‌ازای همه داده های آموزشی آمده‌است. طبق نتایج الگوریتم پیشنهادی، FNE_OGSFS، Group-SAOLA و OGSFS به‌ترتیب به‌طور متوسط روی همه داده‌های آموزشی دارای دقت ۸۰,۱۶٪، ۴۷,۶٪، ۶۲,۶٪ و ۷۱,۳ می‌باشد. این نتایج نشان‌دهنده برتری روش پیشنهادی نسبت به سایر روش‌ها است.

جدول ۲:دقت دسته‌بند SVM به‌ازای الگوریتم‌های انتخاب ویژگی برخط گروهی مختلف

	Proposed_method	FNE_OGSFS	Group-SAOLA	OGSFS
ALLAML	**98.98**	21.72	62.58	70.85
Colon	**87.98**	63.86	62.51	81.62
SMK-CAN-1	**79.02**	34.31	62.51	69.35
credit-g	**72.25**	46.51	69.85	69.54
Sonar	**75.25**	52.93	52.39	69.45
breast.canc	**70.11**	70.5	69.72	70.89
Avg	**80.16**	47.6	62.6	71.3

دقت ویژگی‌های انتخاب‌شده به‌ازای دسته‌بند NB در جدول ۳ نشان آمده‌است. طبق نتایج، الگوریتم پیشنهادی در همه موارد دقت بالاتری نسبت به سایر الگوریتم‌های انتخاب ویژگی برخط گروهی به‌دست آورده‌است. به‌عنوان مثال، دقت الگوریتم پیشنهادی به‌ازای داده آموزشی sonar عبارت است از ۷۰,۷۳ که نشان‌می‌دهد، الگوریتم پیشنهادی تقریباً ۲۴٪، ۲۵٪ و ۸٪ از سایر الگوریتم‌های انتخاب ویژگی گروهی بهتر عمل‌کرده‌است.

جدول ۳: دقت دسته‌بند NB به‌ازای الگوریتم‌های انتخاب ویژگی برخط گروهی مختلف

	Proposed_method	FNE_OGSFS	Group-SAOLA	OGSFS
ALLAML	**96.96**	21.71	53.03	63.65
Colon	**87.94**	56.65	54.6	76.6
SMK-CAN-1	**77.43**	38.37	59.95	59.31
credit-g	**73.55**	46.33	59.18	66.96
Sonar	**70.73**	53.54	52.46	64.59
breast.canc	**71.02**	70.89	70.5	70.63
Avg	**79.7**	47.33	67.8	66.3

در ادامه در جدول ۴ نتایج به‌دست‌آمده به‌ازای دسته‌بند KNN آورده‌شده‌است. نتایج به‌دست‌آمده به‌ازای این دسته‌بند نیز همانند نتایج به‌دست‌آمده به‌ازای دو دسته‌بند قبلی است. همان‌طورکه در جدول ۴ آمده‌است الگوریتم پیشنهادی به‌ازای داده‌های آموزشی ALLAML، colon، SMK-CAN-187 ، credit-g، sonar و breast.cancer بیشترین دقت را به‌دست‌آورده‌است.

جدول ۴: دقت دسته‌بند knn به‌ازای الگوریتم‌های انتخاب ویژگی برخط گروهی مختلف

	Proposed_method	FNE_OGSFS	Group-SAOLA	OGSFS
ALLAML	**98.45**	20.51	64.15	69.68
Colon	**88.01**	57.46	63.47	78.46
SMK-CAN-1	**71.43**	50.56	60.96	66.13
credit-g	**70.03**	20.66	63.14	68.4
Sonar	**76.75**	58.95	51.48	74.66
breast.canc	66.44	69.55	67.35	70.59
Avg	**78.16**	45.66	61.3	70.8

میانگین نتایج حاصل به‌ازای همه داده‌های آموزشی و هر الگوریتم انتخاب ویژگی برخط گروهی در سطر آخر آمده‌است. طبق نتایج، به‌ازای هر سه دسته‌بند دقت به‌دست‌آمده توسط الگوریتم پیشنهادی از سایر الگوریتم‌ها بهتر بوده‌است. همچنین می‌توان‌گفت، OGSFS در رتبه دوم، الگوریتم Group-SAOLA در رتبه سوم و درنهایت الگوریتم FNE-OGSFS در رتبه سوم قراردارد.

۲،۴.ارزیابی روش پیشنهادی از لحاظ تعداد ویژگی‌های انتخابی

در این بخش به مقایسه الگوریتم پیشنهادی و سایر الگوریتم‌های انتخاب ویژگی گروهی برخط بر اساس تعداد ویژگی‌های انتخاب‌شده، می‌پردازیم. نتایج حاصل بعد از ۵ بار اجرای الگوریتم‌های انتخاب ویژگی مختلف در جدول ۵ آمده‌است. همان‌طور که نمودارها نشان‌می‌دهد به‌ازای داده آموزشی ALLAML، الگوریتم پیشنهادی، FNE-OGSFS، Group-SAOLA و OGSFS به‌ترتیب ۲۲، ۲،۵، ۱۱،۶۶ و ۲۴،۳۳ ویژگی را انتخاب‌کرده‌اند. پس الگوریتم FNE-OGSFS نسبت به سایر الگوریتم‌ها تعداد کمتر و OGSFS بیشترین ویژگی را انتخاب‌کرده‌است. نتایج به‌ازای سایر داده‌های آموزشی نیز تقریباً به همین صورت است. به‌طور کلی می‌توان‌گفت الگوریتم FNE-OGSFS معمولاً تعداد کمتر، الگوریتم group-SAOLA دومین کمترین و الگوریتم پیشنهادی سومین کمترین تعداد ویژگی‌ها را انتخاب‌کرده‌است. از این میان، الگوریتم OGSFS معمولاً تعداد بیشتری ویژگی انتخاب‌کرده‌است.

جدول ۵: تعداد ویژگی‌های انتخابی توسط الگوریتم‌های انتخاب ویژگی مختلف

	Proposed_method	FNE_OGSFS	Group-SAOLA	OGSFS
ALLAML	22	**2.5**	11.66	24.33
Colon	5	**1**	4.6	36
SMK-CAN-18	12.66	**1.33**	47.33	59
credit-g	4	**2.33**	3.33	5
Sonar	3.66	**2**	2.66	18.66
breast.cancer	2.33	**1**	2	3.33
Avg	8.27	**1.69**	11.95	24.38

۳،۴. ارزیابی روش پیشنهادی از لحاظ زمان اجرا

متوسط زمان اجرای الگوریتم‌های انتخاب ویژگی برخط گروهی به‌ازای ۵ بار اجرای مستقل در جدول ۶ آمده‌است. طبق نتایج حاصل، متوسط زمان اجرای الگوریتم پیشنهادی معمولاً از سایر الگوریتم‌ها کمتر است. بعد از الگوریتم پیشنهادی، الگوریتم Group-SAOLA زمان اجرای کمتری داشته‌است. در نهایت الگوریتم‌های OGSFS و FNE-OGSFS در رتبه سوم و چهارم قراردارند.

جدول ۶: زمان اجرای الگوریتم‌های انتخاب ویژگی برخط گروهی مختلف

	Proposed_method	FNE_OGSFS	Group-SAOLA	OGSFS
ALLAML	7.0392	12.8143	**0.2743**	1.136
Colon	**0.1866**	2.3559	0.2329	0.4299
SMK-CAN-1	**13.9119**	242.0667	40.3233	124.1812
credit-g	**0.0598**	10.157	0.0668	0.0827
Sonar	**0.0625**	2.1401	0.1231	0.1274
breast.cancer	**0.0406**	0.5778	0.0376	0.089
Avg	**3.3**	44.8	6.76	20.5

۴،۴. پیچیدگی محاسباتی

همان‌طور که گفته‌شد، الگوریتم پیشنهادی با ورود هر گروه از ویژگی‌ها اجرامی‌شود. در ابتدا در فاز اول، میزان ارتباط هر کدام از ویژگی‌های موجود در گروه طبق معادله ۴ محاسبه‌می‌شود. ازآنجاکه این مقدار یک مقدار ثابت است پس می‌توان با m به آن اشاره‌کرد. حال اگر تعداد ویژگی‌های موجود در گروه تازه‌وارد |G| باشد پس می‌توان‌گفت این فاز از مرتبه زمانی $m|G|^2$ است. حال در فاز دوم، میزان افزونگی هر ویژگی با سایر ویژگی‌های دیگر در مجموعه 'S بر اساس رابطه (۵) محاسبه‌می‌شود. اگر پیچیدگی محاسباتی رابطه ۲ را نیز n درنظربگیریم می‌توان‌گفت این فاز نیز از مرتبه زمان $n|S'|^2$ است. پس به‌طور کلی مرتبه اجرایی الگوریتم پیشنهادی عبارت است از

$$O(m|G|^2+n|S'|^2)\ \square\ O(\max(|G|^2,|S'|^2))$$

می‌توان‌گفت، الگوریتم پیشنهادی در یک زمان خطی که وابسته به تعداد ویژگی‌های انتخاب‌شده تا آن لحظه یا اندازه گروه است به فرایند انتخاب ویژگی می‌پردازد. بنابراین ازآنجاکه مرتبه زمانی انتخاب ویژگی‌ها در الگوریتم پیشنهادی به‌صورت خطی است می‌توان‌گفت الگوریتم پیشنهادی مقیاس‌پذیر است.

۵. نتیجه‌گیری و کارهای آینده

یکی از چالش‌های اساسی که این روزها با آن سروکار داریم جریانی‌بودن داده‌های آموزشی است که باعث بروز مشکلات زیادی شده‌است. به این معنی که داده‌های آموزشی به‌صورت تدریجی و با مرور زمان به مجموعه اضافه‌شده و همین امر باعث افزایش بی‌رویه داده‌های آموزشی شده‌است. از طرف دیگر در این داده‌ها معمولاً مسأله ازدحام ابعاد وجوددارد به این معنی که که داده‌های آموزشی تحت تأثیر تعداد بالای ویژگی‌ها قرارگرفته و باعث‌می‌شوند الگوریتم‌های یادگیری ماشین با مشکل روبرو شوند. این امر به این دلیل است که تعداد زیادی از این ویژگی‌ها نامربوط و یا اضافی هستند. تاکنون الگوریتم‌های انتخاب ویژگی زیادی در جریان داده‌های گروهی ویژگی‌ها ارائه‌شده‌است. اما هیچ‌کدام از آن‌ها نتواسته‌است همه چالش‌های مربوط به این دسته از داده‌ها را برطرف‌کند. به همین منظور، در این مقاله یک راهکار انتخاب ویژگی در جریان داده‌های گروهی ویژگی ارائه‌شد. راهکار پیشنهادی با استفاده از اطلاعات متقابل طی دو فاز، ویژگی‌های غالب را شناسایی و ویژگی‌های اضافی را حذف‌می‌کند. در فاز اول ویژگی‌های مرتبط بر اساس اطلاعات متقابل انتخاب و در فاز دوم بر اساس اطلاعات متقابل ویژگی‌های اضافی حذف‌می‌شوند. سپس روش پیشنهادی در مقایسه با سه الگوریتم انتخاب ویژگی برخط گروهی و ۶ داده آموزشی معروف و پرکاربرد با استفاده از چندین معیار مختلف مانند تعداد ویژگی‌های انتخاب‌شده، زمان اجرا و دقت انواع دسته‌بندها موردارزیابی قرارگرفت. نتایج نشان‌داد الگوریتم پیشنهادی معمولاً تعداد کمتری ویژگی و در زمان کمتری با دقت دسته بندی بالا انتخاب‌می‌کند. همچنین الگوریتم پیشنهادی نسبت به سایر الگوریتم‌ها مقیاس‌پذیرتر است. در آینده قصد داریم، الگوریتم پیشنهادی را به‌گونه‌ایی تغییردهیم که در داده‌های آموزشی چندبرچسبی نیز به‌خوبی بتواند کارایی خود را حفظکند.

References

[۱] Z. Fang, J.-N. Hwang, X. Huo, H.-J. Lee, and J. Denzler, "Emergent Techniques and Applications for Big Visual Data," *International Journal of Digital Multimedia Broadcasting*, vol. 2017, p. 6468502, 2017/10/31 2017, doi: 10.1155/2017/6468502.

[۲] T. Zhang and B. Yang, "Big data dimension reduction using PCA," in *2016 IEEE international conference on smart cloud (SmartCloud)*, 2016: IEEE, pp. 152-157.

[۳] M. Hariri and H. Najafy, "Improve the Quality of Mammogram Images by Image Processing Techniques," *Intelligent Multimedia Processing and Communication Systems (IMPCS)*, vol. 3, no. 1, pp. 57-69, 2022. [Online]. Available: https://impcs.zanjan.iau.ir/article_696998_4bc75e4c5155a48f9fb7b94ac977a892.pdf.

[۴] M. Köppen, "The curse of dimensionality," in *5th online world conference on soft computing in industrial applications (WSC5)*, 2000, vol. 1, pp. 4-8.

[۵] M. Najafi, M. Afzali, and M. Moradi, "Use data mining to identify factors affecting students' academic failure," *Intelligent Multimedia Processing and Communication Systems (IMPCS)*, vol. 2, no. 1, pp. 23-33, 2021. [Online]. Available: https://impcs.zanjan.iau.ir/article_682093_7de40314874a6374a5cdc7b20b69a455.pdf.

[۶] A. Jović, K. Brkić, and N. Bogunović, "A review of feature selection methods with applications," in *2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO)*, 2015: Ieee, pp. 1200-1205.

[۷] L. Sun, S. Fu, and F. Wang, "Decision tree SVM model with Fisher feature selection for speech emotion recognition," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2019, no. 1, pp. 1-14, 2019.

[۸] Y. Liu, F. Tang, and Z. Zeng, "Feature selection based on dependency margin," *IEEE Transactions on Cybernetics*, vol. 45, no. 6, pp.۲۰۱۴ ,۱۲۲۱-۱۲۰۹ .

[۹] S. Ahmed, Y. Lee, S.-H. Hyun, and I. Koo, "Covert cyber assault detection in smart grid networks utilizing feature selection and euclidean distance-based machine learning," *Applied Sciences*, vol. 8, no. 5, p. 772, 2018.

[۱۰] J. R. Vergara and P. A. Estévez, "A review of feature selection methods based on mutual information," *Neural computing and applications*, vol. 24, pp. 175-186, 2014.

[۱۱] S. Maldonado and R. Weber, "A wrapper method for feature selection using support vector machines," *Information Sciences*, vol. 179, no. 13, pp. 2208-2217, 2009.

[۱۲] H. Liu, M. Zhou, and Q. Liu, "An embedded feature selection method for imbalanced data classification," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 3, pp. 703-715, 2019.

[۱۳] T. N. Lal, O. Chapelle, J. Weston, and A. Elisseeff, "Embedded methods," *Feature Extraction: Foundations and Applications*, pp. 137-165, 2006.

[۱۴] W. Ding *et al.*, "Subkilometer crater discovery with boosting and transfer learning," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 4, pp. 1-22, 2011.

[۱۵] X. Wu, K. Yu, W. Ding, H. Wang, and X. Zhu, "Online feature selection with streaming features," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 5, pp..۲۰۱۲ ,۱۱۹۲-۱۱۷۸

[۱۶] D. You *et al.*, "Online feature selection for multi-source streaming features," *Information Sciences*, vol. 590, pp. 267-295, 2022/04/01/ 2022, doi: https://doi.org/10.1016/j.ins.2022.01.008.

[۱۷] D. Paul, A. Jain, S. Saha, and J. Mathew, "Multi-objective PSO based online feature selection for multi-label classification," *Knowledge-Based Systems*, vol. 222, p. 106966, 2021.

[۱۸] S. Eskandari and M. M. Javidi, "Online streaming feature selection using rough sets," *International*

Journal of Approximate Reasoning, vol. 69, pp. 35-57, 2016.

[۱۹] J. Zhou, D. Foster, R. Stine, and L. Ungar, "Streaming feature selection using alpha-investing," in *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, 2005, pp. 384-393.

[۲۰] S. Perkins, K. Lacker, and J. Theiler, "Grafting: Fast, incremental feature selection by gradient descent in function space," *The Journal of Machine Learning Research*, vol. 3, pp. 1333-1356, 2003.

[۲۱] M. Wang, H. Li, D. Tao, K. Lu, and X. Wu, "Multimodal Graph-Based Reranking for Web Image Search," *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4649-4661, 2012, doi: 10.1109/TIP.2012.2207397.

[۲۲] P. Zhou, N. Wang, and S. Zhao, "Online group streaming feature selection considering feature interaction," *Knowledge-Based Systems*, vol. 226, p. 107157, 2021/08/17/ 2021, doi: https://doi.org/10.1016/j.knosys.2021.107157.

[۲۳] J. Wang *et al.*, "Online feature selection with group structure analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 11, pp. 3029-3041, 2015.

[۲۴] K. Yu, X. Wu, W. Ding, and J. Pei, "Scalable and accurate online feature selection for big data," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 11, no. 2, pp. 1-3.۲۰۱۶ ,۹

[۲۵] T. Dokeroglu, A. Deniz, and H. E. Kiziloz, "A comprehensive survey on recent metaheuristics for feature selection," *Neurocomputing*, vol. 494, pp. 269-296, 2022/07/14/ 2022, doi: https://doi.org/10.1016/j.neucom.2022.04.083.

[۲۶] U. M. Khaire and R. Dhanalakshmi, "Stability of feature selection algorithm: A review," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 4, pp. 1060-1073, 2022/04/01/ 2022, doi: https://doi.org/10.1016/j.jksuci.2019.06.012.

[۲۷] M. Vahmiyan, M. Kheirabadi, and E. Akbari, "Feature selection methods in microarray gene expression data: a systematic mapping study," *Neural Comput. Appl.*, vol. 34, no. 22, pp. 19675–19702, 2022, doi: 10.1007/s00521-022-07661-z.

[۲۸] M. H. Nadimi-Shahraki, H. Zamani, and S. Mirjalili, "Enhanced whale optimization algorithm for medical feature selection: A COVID-19 case study," *Computers in Biology and Medicine*, vol. 148, p. 105858, 2022/09/01/ 2022, doi: https://doi.org/10.1016/j.compbiomed.2022.105858.

[۲۹] M. Zivkovic, C. Stoean, A. Chhabra, N. Budimirovic, A. Petrovic, and N. Bacanin, "Novel Improved Salp Swarm Algorithm: An Application for Feature Selection," *Sensors*, vol. 22, no. 5, p. 1711, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/5/.۱۷۱۱

[۳۰] M. M. Javidi and S. Eskandari, "Streamwise feature selection: a rough set method," *International Journal of Machine Learning and Cybernetics*, vol. 9, pp. 667-676, 2018.

[۳۱] J. Wang, P. Zhao, S. C. Hoi, and R. Jin, "Online feature selection and its applications," *IEEE Transactions on knowledge and data engineering*, vol. 26, no. 3, pp. 698-710, 2013.

[۳۲] K. Yu, X. Wu, W. Ding, and J. Pei, "Towards scalable and accurate online feature selection for big data," in *2014 IEEE International Conference on Data Mining*, 2014: IEEE, pp. 660-669.

[۳۳] M. Rahmaninia and P. Moradi, "OSFSMI: Online stream feature selection method based on mutual information," *Applied Soft Computing*, vol. 68, pp. 733-746, 2018/07/01/ 2018, doi: https://doi.org/10.1016/j.asoc.۲۰۱۷/۰۸/۰۳۴.

[۳۴] H. Li, X. Wu, Z. Li, and W. Ding, "Group feature selection with streaming features," in *2013 IEEE 13th International Conference on Data Mining*, 2013: IEEE, pp. 1109-1114.

[۳۵] J. Xu, Y. Sun, K. Qu, X. Meng, and Q. Hou, "Online group streaming feature selection using entropy-based uncertainty measures for fuzzy neighborhood rough sets," *Complex & Intelligent Systems*, vol. 8, no. 6, pp. 5309-5328, 2022.

[۳۶] E. Fix and J. L. Hodges, "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties," *International Statistical Review / Revue Internationale de Statistique*, vol. 57, no. 3, pp. 238-247, 1989, doi: 10.2307/1403797.

[۳۷] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121-167, 1998/06/01 1998, doi: 10.1023/A:1009715923555.

[۳۸] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2.۰۰۹