

## پیش بینی ابعاد شخصیتی یادگیرندگان در محیط یادگیری الکترونیکی با استفاده از ماشین بردار پشتیبان دو قلو کمترین مربعات

امیر محمود میر<sup>۱</sup>، سمیه فتاحی<sup>۲\*</sup>، جلال الدین نصیری<sup>۳</sup>

۱- گروه هوش مصنوعی، دانشکده برق و کامپیوتر، واحد تهران شمال، دانشگاه آزاد اسلامی، تهران، ایران

۲- استادیار، پژوهشگاه علوم و فناوری اطلاعات ایران، تهران، ایران

۳- استادیار، پژوهشگاه علوم و فناوری اطلاعات ایران، تهران، ایران

\*نویسنده مسئول: fatahi@irandoc.ac.ir

### خلاصه

یکی از اهداف مهم در محیط های آموزشی، بهبود عملکرد یادگیرندگان است. برای تحقق این هدف، پژوهش های زیادی صورت گرفته است که بخشی از آنها مربوط به شناسایی تیپ شخصیتی یادگیرندگان است. با آگاهی از تیپ شخصیتی یادگیرنده، می توان آموزش را متناسب با ویژگی های یادگیرنده طراحی کرد. در این مقاله، مسئله اصلی پیش-بینی ابعاد شخصیتی یادگیرندگان در یک محیط آموزش الکترونیکی است. برای پیش بینی ابعاد شخصیتی یادگیرندگان، داده هایی از تعاملات یادگیرندگان با یک محیط آموزش الکترونیکی جمع آوری شد و ۱۹ ویژگی از این تعاملات استخراج شده است. با استفاده از ویژگی های مستخرج از تعاملات و طبقه بندی کننده ماشین بردار پشتیبان دو قلو کمترین مربعات، مدلی قابل اطمینان برای پیش بینی ابعاد شخصیتی یادگیرندگان در یک محیط آموزش الکترونیکی بدست آمده است. نتایج نشان می دهد که با دقت مناسبی می توان شخصیت یادگیرندگان را پیش بینی نمود.

کلمات کلیدی: شخصیت، یادگیری الکترونیکی، ماشین بردار پشتیبان دو قلو کمترین مربعات، MBTI

### ۱. مقدمه

دسترسی به اینترنت و کامپیوتر فرصت هایی را برای آموزش الکترونیکی<sup>۱</sup> ایجاد کرده اند، از مزایای آموزش الکترونیکی می توان به دسترسی راحت تر به منابع و همچنین آزادی عمل کاربران جهت زمان بندی اشاره کرد. با این حال، یک نقطه ضعف مهم آموزش های الکترونیکی، نداشتن جذابیت و پویایی است. همین نقطه ضعف، انگیزه محققان را برای تشخیص یا پیش بینی ابعاد شخصیتی دانشجویان در آموزش الکترونیکی به منظور بهبود این سیستم ها افزایش داده است [۱]. تشخیص خودکار شخصیت یادگیرندگان کمک می کند تا آموزش براساس روشی مناسب هر فرد انجام شود و سیستم های آموزش های الکترونیکی، برای کاربران جذاب تر باشند و کارایی کاربران افزایش یابد.

در روانشناسی، شخصیت به عنوان یک موضوع تحقیقاتی مهم و تاثیرگذار است. به دلیل اینکه شخصیت نشان دهنده عوامل مهمی مانند سلامت روحی و جسمی، کیفیت روابط اجتماعی، کارایی فرد در محیط کار یا محیط آموزشی است [۲]. همچنین تحقیقاتی بر روی تاثیر ابعاد شخصیتی روی عملکرد فرد در شغلش انجام گرفته است که اهمیت شناخت

<sup>1</sup> E-Learning

افراد بر اساس ابعاد شخصیتی شان را نشان می دهد [۳]. بطوریکه در فرآیند استخدام افراد، برای تشخیص تیپ شخصیتی آن ها آزمون گرفته می شود تا آن ها در شغلی متناسب با ویژگی های شخصیتی شان قرار داده شوند.

تشخیص یا پیش بینی ابعاد شخصیتی افراد نه تنها در محیط کار، بلکه در محیط آموزشی (حضور و یا الکترونیکی) نیز حائز اهمیت است. با پیش بینی ابعاد شخصیتی دانشجویان در محیط آموزشی، می توان طبق تیپ شخصیتی شان، شیوه مناسب آموزش را برای آنها در نظر گرفت. تحقیقاتی بر روی تاثیر تیپ شخصیتی بر روی عملکرد درسی دانشجویان انجام گرفته است و نتایج نشان می دهد که آگاهی از تیپ شخصیتی یادگیرندگان، اساتید را در جهت طراحی روش آموزش مناسب برای یادگیرندگان کمک می کند [۴].

در این مقاله، مسئله مورد مطالعه پیش بینی ابعاد شخصیتی یادگیرندگان در یک محیط آموزشی الکترونیکی است. در تحقیقات پیشین، برای تشخیص تیپ شخصیتی یادگیرنده از روش های ماشین بردار پشتیبان<sup>۱</sup>، شبکه های عصبی، K-means و یادگیری عمیق<sup>۲</sup> استفاده شده است [۵] [۶] [۷] [۲]. این مقاله از دو منظر با تحقیقات پیشین متفاوت است:

۱. برای پیش بینی ابعاد شخصیتی یادگیرندگان، یک مجموعه داده جدید از یک محیط آموزش الکترونیکی جمع آوری شده است.

۲. از طبقه بندی کننده ماشین بردار پشتیبان دو قلو کمترین مربعات [۸] برای پیش بینی ابعاد شخصیتی یادگیرندگان استفاده شده است.

نتایج نشان می دهد که با استفاده از ویژگی های مستخرج و طبقه بندی کننده ماشین بردار پشتیبان دو قلو کمترین مربعات، مدلی قابل اطمینان برای پیش بینی ابعاد شخصیتی یادگیرندگان در یک محیط آموزش الکترونیکی بدست آمده است.

## ۲. مدل شخصیتی MBTI

در روانشناسی تعاریف متعددی از شخصیت وجود دارد. شولتز شخصیت را جنبه های درونی و بیرونی یک فرد تعریف می کند که رفتار انسان را در حالت های مختلف تحت تاثیر قرار می دهد [۹]. کارل یونگ<sup>۳</sup> مدل شخصیتی MBTI<sup>۴</sup> را معرفی کرد. از این مدل شخصیتی برای تعیین سبک یادگیری متناسب هر فرد نیز استفاده می شود [۷].

این پژوهش نیز از مدل شخصیتی MBTI برای پیش بینی ابعاد شخصیتی یادگیرندگان استفاده می کند. این مدل دارای چهار بعد شخصیتی است که شامل برونگرا - درونگرا<sup>۵</sup>، حسی - شهودی<sup>۶</sup>، فکری - احساسی<sup>۷</sup> و قضاوتی - ادراکی<sup>۸</sup> می باشد. هر فرد بر اساس مدل MBTI چهار بعد شخصیتی دارد. برای مثال یک یادگیرنده می تواند برونگرا، حسی، فکری و قضاوتی باشد. شکل ۱ چهار بعد شخصیتی را در مدل MBTI نشان می دهد.

<sup>1</sup> Support Vector Machine

<sup>2</sup> Deep learning

<sup>3</sup> Carl Jung

<sup>4</sup> The Myers-Briggs Type Indicator

<sup>5</sup> Extrovert-Introvert

<sup>6</sup> Sensing-iNtuition

<sup>7</sup> Thinking-Feeling

<sup>8</sup> Judging-Perceiving



شکل ۲ - ۱۶ تیپ شخصیتی در مدل MBTI

شکل ۱ - چهار بعد شخصیتی مدل MBTI

چهار بعد شخصیتی ذکر شده در شکل ۱ در مجموع ۱۶ تیپ شخصیتی منحصر به فرد را در مدل MBTI می‌سازد. شکل ۲ تیپ‌های شخصیتی در مدل MBTI را نشان می‌دهد. برای مثال، تیپ شخصیتی ISTJ، دارای ابعاد شخصیتی درونگرا، حسی، فکری و قضاوتی است.

### ۳. تحقیقات پیشین

تحقیقات زیادی بر روی شخصیت در علوم کامپیوتر و تعامل انسان و رایانه صورت گرفته است. برخی از آن‌ها در اینجا آورده شده است: فلدر<sup>۱</sup> و همکاران [۴]، تاثیر تیپ شخصیتی بر روی عملکرد دانشجویان را بررسی کردند. در پژوهش آنان، از مدل MBTI برای تشخیص تیپ شخصیتی دانشجویان رشته مهندسی شیمی در دانشگاه ایالتی کارولینای شمالی<sup>۲</sup> استفاده شده است. نتایج این پژوهش نشان می‌دهد که مدل MBTI یک ابزار مفید برای کمک کردن به اساتید رشته-های مهندسی جهت طراحی یک شیوه تدریس مناسب است که باعث بهبود عملکرد تمام دانشجویان گردد. یافته‌های این تحقیق نشان می‌دهد که دانشجویان با بعد شخصیتی شهودی نسبت به دانشجویان حسی، مفاهیم انتزاعی را بهتر درک می‌کنند.

گانگ<sup>۳</sup> و ونگ<sup>۴</sup> [۵]، کاربرد روش ماشین بردار پشتیبان در تشخیص شخصیت در محیط یادگیری الکترونیکی را بررسی کردند. در این پژوهش، مدل فلدر-سیلورمن<sup>۵</sup> برای مسئله تشخیص تیپ شخصیتی یادگیرنده پیشنهاد شده است. همچنین یک مدل پویا پیشنهاد شده است که قادر به تشخیص تیپ شخصیتی یادگیرنده و فراهم کردن محتوای آموزشی مناسب یادگیرنده می‌باشد و ماشین بردار پشتیبان، طبقه‌بندی کننده مدل پیشنهادی می‌باشد.

<sup>1</sup> Felder

<sup>2</sup> North Carolina State University

<sup>3</sup> Gong

<sup>4</sup> Wang

<sup>5</sup> Felder-Sliverman

کلهتگی<sup>۱</sup> و همکاران [۶] اقدام به تشخیص تیپ شخصیتی کاربران از توییت های آنان در شبکه اجتماعی Twitter کردند. در روش پیشنهادی این پژوهش، ویژگی هایی از توییت ها بر اساس اطلاعات مربوط به گرامر و رفتار اجتماعی استخراج شده است. برای تشخیص تیپ شخصیتی کاربران، از طبقه بندی کننده از نوع شبکه عصبی پرسپترون چند لایه استفاده شده است. در این پژوهش کاربردهای تشخیص تیپ شخصیتی کاربران در شبکه های اجتماعی ذکر شده است، از جمله پیشنهاد دادن کتاب یا موسیقی های مورد علاقه کاربران.

فتاحی و همکاران [۷]، مسئله پیش بینی ابعاد شخصیتی دانشجویان در یک محیط آموزش الکترونیکی را بررسی کرده اند. این پژوهش از مدل MBTI برای مسئله پیش بینی ابعاد شخصیتی دانشجویان استفاده می کند. ابتدا از یک محیط آموزش الکترونیکی واقعی داده های تعاملات کاربران با محیط یادگیری الکترونیکی جمع آوری شده و سپس با روش خوشه بندی K-means ویژگی هایی که تفکیک بهتری برای مسئله پیش بینی ابعاد شخصیتی دانشجویان دارند، انتخاب شده اند. بهترین نتیجه بدست آمده برای پیش بینی بعد شخصیتی فکری - احساسی است که دقت آن ۷۲ درصد می باشد.

ما<sup>۲</sup> و لیو<sup>۳</sup> [۲]، از یادگیری عمیق از نوع شبکه عصبی بازگشتی برای پیش بینی تیپ شخصیتی افراد از متن نوشته شده توسط آن ها استفاده کردند. در روش پیشنهادی این تحقیق، از متون پنج جمله ای برای آموزش طبقه بندی کننده استفاده شده است. در این پژوهش، پیش بینی تیپ شخصیتی افراد بر اساس مدل MBTI است. بیشترین دقت بدست آمده ۳۷ درصد می باشد که نشان می دهد که تشخیص تیپ شخصیتی افراد از متن، کار آسانی نیست.

#### ۴. مجموعه داده

#### ۴-۱ جمع آوری و آماده سازی داده ها

برای انجام این پژوهش، از سامانه آموزش الکترونیکی بر پایه مدل<sup>۴</sup> که در دانشگاه تهران در حال اجراست، استفاده شده است و کاربران جهت یادگیری مباحث درسی در این سامانه وارد می شوند. در سامانه آموزش الکترونیکی مذکور، ۲۰۲ دانشجو مهندسی برق و کامپیوتر از دانشگاه تهران عضو شده و در واحد درسی «مبانی کامپیوتر و برنامه نویسی» ثبت نام کرده اند. جمع آوری داده ها دو فاز دارد. در فاز اول ابتدا دانشجویان پرسش نامه مدل MBTI را جواب می دهند تا تیپ شخصیتی آن ها مشخص شود.

در فاز دوم، یک مبحث به دانشجو آموزش داده می شود که فرآیند آن در این سامانه به صورت زیر است:

۱. ابتدا کاربر به ۵ سوال درباره مبحث «آرایه ها و اشاره گرها» پاسخ می دهد. برای هر سوال از کاربر سوالات

زیر پرسیده می شود:

۱. آیا نیازی به راهنمایی معلم دارند یا خیر.
۲. آیا راهنمایی معلم برای آن ها مفید بوده است یا خیر.
۳. میزان تلاش کاربر برای پاسخ دادن به هر سوال پرسیده می شود.

<sup>1</sup> Kalghatgi

<sup>2</sup> Ma

<sup>3</sup> Liu

<sup>4</sup> <https://moodle.org/>

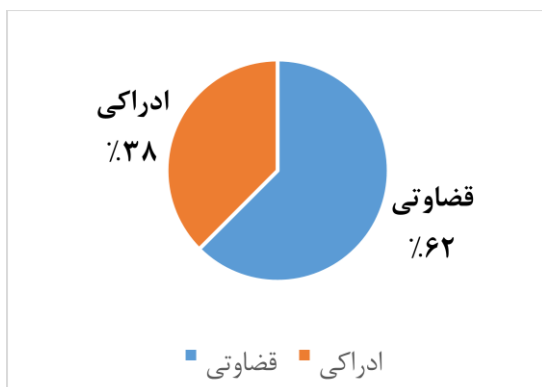
۲. پس از اتمام مرحله ۱، آموزشی متناسب با سطح دانش دانشجو درباره میحث «آرایه‌ها و اشاره‌گرها» به وی داده می‌شود. پس از آموزش، از کاربر ۵ سوال دیگر پرسیده می‌شود. مانند مرحله ۱، سه سوال مذکور برای هر سوال از دانشجو پرسیده می‌شود.

۳. پس از اتمام مرحله ۱ و ۲، میزان خوشایندی دانشجو از سوالات و نحوه آموزش پرسیده می‌شود.

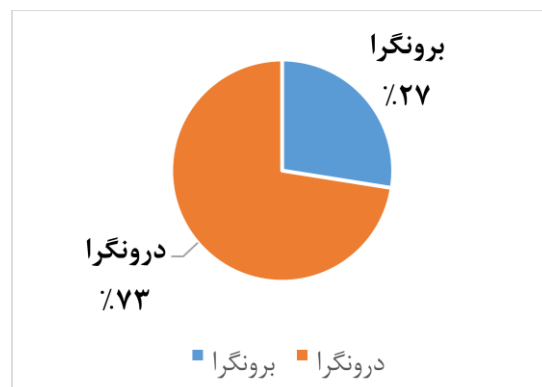
پس از اتمام فرآیند آموزش کاربران در این سامانه، مجموعه داده‌ای برای انجام این پژوهش ایجاد شده است. برای پیش‌بینی ابعاد شخصیتی دانشجویان، از این مجموعه داده ویژگی‌هایی استخراج می‌شود که در بخش ۱-۵ توضیح داده شده است.

#### ۲-۴ توزیع مجموعه داده‌ها

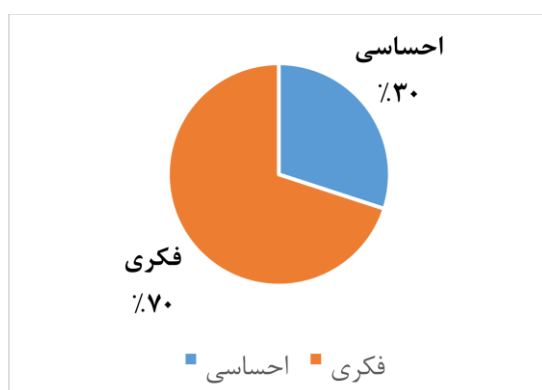
داده جمع آوری شده که در بخش ۱-۴ توضیح داده شد، به ۴ مجموعه داده تقسیم می‌شود. از آنجایی که در این پژوهش می‌خواهیم چهار بعد شخصیتی مدل MBTI را پیش‌بینی کنیم، برای هر بعد شخصیتی نیاز به یک مجموعه داده مجزا داریم که طبقه‌بندی کننده را بر اساس آن آموزش دهیم. مجموعه داده‌ها در این پژوهش به طور کلی شامل قضاوتی-ادراکی، درونگرا-برونگرا، فکری-احساسی، حسدی-شهودی می باشد که توزیع شان به ترتیب در شکل ۳، ۴، ۵ و ۶



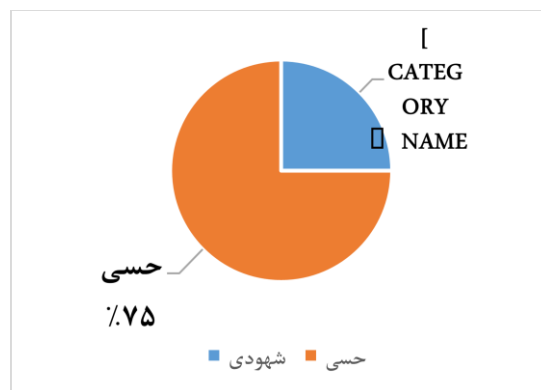
شکل ۳ - توزیع مجموعه داده قضاوتی و ادراکی



شکل ۴ - توزیع مجموعه داده برونگرا و درونگرا



شکل ۵ - توزیع مجموعه داده فکری و احساسی



شکل ۶ - توزیع مجموعه داده حسدی و شهودی

۶ آمده است.

همانطور که در شکل های ۳ تا ۶ مشخص است، چهار مجموعه داده بدست آمده غیر متوازن<sup>۱</sup> هستند. به عبارت دیگر، تعداد نمونه های کلاس ها نسبت به هم تفاوت مشهودی دارند. به عنوان نمونه، مجموعه داده حسی - شهودی شدیداً نامتوازن است، بطوریکه که از کل نمونه ها فقط ۲۵ درصد شهودی است. نامتوازن بودن این مجموعه داده ها باعث می شود که بیشتر الگوریتم های یادگیری به سمت یک کلاس گرایش پیدا کنند. همین موضوع باعث کاهش شدید دقت و عملکرد بیشتر الگوریتم یادگیری می شود و از تشخیص یک کلاس ناتوان می ماند. برای غلبه بر این مشکل، از روش ماشین بردار پشتیبان دو قلو کمترین مربعات به عنوان طبقه بندی کننده برای پیش بینی ابعاد شخصیتی دانشجویان استفاده شده است. از ویژگی های روش LS-TSVM می توان به عملکرد خوب روی مجموعه داده های نامتوازن اشاره کرد.

## ۵. روش پیشنهادی این پژوهش

### ۵-۱ ویژگی ها

ویژگی های استخراج شده برای پیش بینی ابعاد شخصیتی دانشجویان در جدول ۱ ذکر شده است. تعداد ویژگی های استخراج شده برابر با ۱۹ می باشد. ویژگی های ۹ تا ۱۳ مربوط به قبل از آموزش یادگیرنده و ویژگی های ۱۴ تا ۱۸ مربوط به بعد از آموزش یادگیرنده می باشد. نحوه آموزش یادگیرنده در بخش ۱-۴ توضیح داده شده است.

جدول ۱- ویژگی های استخراج شده از داده ها

ویژگی ها		
۱- درصد درونگرایی	۸- درصد ادراکی	۱۵- درصد راهنمایی از معلم
۲- درصد برونگرایی	۹- درصد پاسخ صحیح به سوالات	۱۶- درصد مفید بودن راهنمایی معلم
۳- درصد حسی	۱۰- درصد راهنمایی از معلم	۱۷- درصد میزان تلاش برای پاسخ دادن
۴- درصد شهودی	۱۱- درصد مفید بودن راهنمایی معلم	۱۸- درصد اتمام فعالیت محول شده
۵- درصد فکری	۱۲- درصد میزان تلاش برای پاسخ دادن	۱۹- میزان خوشایندی
۶- درصد احساسی	۱۳- درصد اتمام فعالیت محول شده	
۷- درصد قضاوتی	۱۴- درصد پاسخ صحیح به سوالات	

### ۵-۲ ماشین بردار پشتیبان دو قلو کمترین مربعات

ماشین بردار پشتیبان یا SVM [۱۰]، یک روش یادگیری است که در مسائل مختلف کارایی خوبی از خود نشان داده است. در ماشین بردار پشتیبان به دنبال یک ابرصفحه جدا کننده با بیشترین حاشیه هستیم. ماشین بردار پشتیبان نظر پژوهشگران را به خود جلب کرد و در سال های اخیر توسعه های زیادی از آن مطرح شد که یکی از آن ها ماشین بردار پشتیبان دو قلو<sup>۲</sup> [۱۱] یا Twin-SVM می باشد که نسبت به SVM استاندارد سرعت بیشتری دارد.

<sup>1</sup> Unbalanced

<sup>2</sup> Twin Support Vector Machine



ماشین بردار پشتیبان دوقلو برخلاف SVM استاندارد، به دنبال دو ابرصفحه غیر موازی می باشد، بطوریکه روی هر کلاس یک ابرصفحه غیر موازی وجود دارد. برای مثال روی هر کلاس مجموعه داده حسی- شهودی، یک ابرصفحه غیر موازی وجود دارد. ماشین بردار پشتیبان دوقلو از SVM استاندارد حدوداً ۴ برابر سریعتر است. با این حال برای رسیدن به دو ابرصفحه غیر موازی نیاز به حل کردن دو مسئله بهینه سازی کوچک می باشد. حل این دو مسئله بهینه سازی می تواند بسته به ابعاد بردار ویژگی زمان بر باشد. به همین دلیل در این پژوهش از نسخه بهبود یافته Twin-SVM یعنی LS-TSVM به عنوان طبقه بندی کننده برای پیش بینی ابعاد شخصیتی دانشجویان استفاده شده است.

ماشین بردار پشتیبان دوقلو کمترین مربعات یا LS-TSVM نیز مانند Twin-SVM به دنبال دو ابرصفحه غیر موازی است. بطوریکه روی داده های هر کلاس یک ابر صفحه غیر موازی قرار می گیرد. تفاوت اصلی ماشین بردار پشتیبان دو قلو کمترین مربعات با Twin-SVM نحوه بدست آوردن این دو ابرصفحه می باشد که دو مسئله بهینه سازی با قید تعریف می شود که در رابطه ی ۱ و ۲ نشان داده شده است.

$$\text{Min}_{w^{(1)}, b^{(1)}} \frac{1}{2} (Aw^{(1)} + eb^{(1)})^T (Aw^{(1)} + eb^{(1)}) + \frac{C1}{2} y^T y \quad (1)$$

$$\text{s.t. } -(Bw^{(1)} + eb^{(1)}) + y = e$$

$$\text{Min}_{w^{(2)}, b^{(2)}} \frac{1}{2} (Bw^{(2)} + eb^{(2)})^T (Bw^{(2)} + eb^{(2)}) + \frac{C2}{2} y^T y \quad (2)$$

$$\text{s.t. } (Aw^{(2)} + eb^{(2)}) + y = e$$

بردار  $w(i)$  مختصات ابرصفحه  $i$  می باشد و  $b$  نیز بایاس است. ماتریس  $A$  بیانگر داده های کلاس ۱ و ماتریس  $B$  نیز بیانگر داده های کلاس ۱- است. بردار  $e$  با ابعاد مناسب مقادیر یک را در خود دارد. متغیر  $y$ ، متغیر لغزش می باشد. در رابطه ی ۱ و ۲ دو پارامتر خطای  $C1$  و  $C2$  وجود دارد که برای رسیدن به کمترین خطا نیاز به تنظیم شدن دارد که در بخش پیاده سازی دامنه آنها توضیح داده شده است.

به طور کلی رابطه ۱ بیان می کند که دو مقدار  $w(1)$  و  $b(1)$  باید طوری انتخاب شود که داده های کلاس ۱ (ماتریس  $A$ ) کمترین فاصله را با ابرصفحه غیر موازی ۱ داشته باشد و همچنین قید رابطه ی ۲ بیانگر فاصله ی داده های کلاس ۱- (ماتریس  $B$ ) از ابرصفحه غیر موازی ۱ است. رابطه ی ۲ نیز مشابه رابطه ی ۱ تفسیر می شود با این تفاوت که ماتریس  $B$  به ابرصفحه غیر موازی ۲ نزدیک تر است و از ابرصفحه غیر موازی ۱ دورتر می باشد. از آنجایی که قید هر دو رابطه ی ۱ و ۲ مساوی با ماتریس  $e$  است، می توان دو مسئله بهینه سازی فوق را با دستگاه معادلات خطی حل نمود که همین موضوع زمان آموزش LS-TSVM را بسیار کمتر از Twin-SVM می کند. رابطه ی ۳ و ۴ راه حل دو مسئله بهینه سازی ذکر شده در رابطه ی ۱ و ۲ هستند.

$$\begin{bmatrix} w^{(1)} \\ b^{(1)} \end{bmatrix} = - (F^T F + \frac{1}{C1} E^T E)^{-1} F^T e \quad (3)$$

$$\begin{bmatrix} w^{(2)} \\ b^{(2)} \end{bmatrix} = (E^T E + \frac{1}{C2} F^T F)^{-1} E^T e \quad (4)$$

ماتریس E برابر است با [A e] و ماتریس F نیز برابر با [B e] می باشد. اکنون با رابطه ۳ و ۴ می توان مختصات دو ابرصفحه غیر موازی که هر کدام روی یک کلاس قرار می گیرند را بدست آورد که این دو ابرصفحه غیر موازی در رابطه ی ۵ تعریف شده است.

$$x^T w^{(1)} + b^{(1)} = 0 \quad x^T w^{(2)} + b^{(2)} = 0 \quad (5)$$

برای فهمیدن اینکه داده ارزیابی به کدام کلاس تعلق دارد، فاصله عمودی<sup>۱</sup> آن داده را از دو ابرصفحه غیر موازی محاسبه می کنیم و سپس داده ارزیابی به کلاسی تعلق دارد که فاصله آن با ابرصفحه غیر موازی آن کلاس کمتر باشد. رابطه ی ۶ نشان دهنده فاصله عمودی داده ارزیابی از کلاس ۱ و رابطه ی ۷ نشان دهنده فاصله ی عمودی داده ارزیابی از کلاس ۱- می باشد.

$$|x^T w^{(1)} + b^{(1)}| \quad (6)$$

$$|x^T w^{(2)} + b^{(2)}| \quad (7)$$

توضیحات و رابطه های آورده شده برای حل مسائل خطی توسط ماشین بردار پشتیبان دوقلو کمترین مربعات است. با این حال LS-TSVM برای حل مسائل غیر خطی نیز استفاده می شود که در آن از یک تابع هسته<sup>۲</sup> می شود که توضیحات و رابطه های نسخه غیر خطی این روش یادگیری در مقاله اصلی [۸] ذکر شده است.

### ۳-۵ معیارهای ارزیابی روش پیشنهادی

در این بخش معیارهای ارزیابی برای سنجش عملکرد روش پیشنهادی این مقاله که هدف آن پیش بینی ابعاد شخصیتی یادگیرندگان می باشد، توضیح داده شده است. برای سنجش عملکرد روش پیشنهادی از معیار های دقت<sup>۳</sup>، بازخوانی<sup>۴</sup>، صحت<sup>۵</sup> و F1 استفاده شده که این معیار ها به ترتیب در رابطه ی ۸، ۹، ۱۰ و ۱۱ نشان داده شده است. معیار دقت به تنهایی برای سنجش عملکرد روش پیشنهادی کافی نمی باشد. به دلیل اینکه مجموعه داده این تحقیق نامتوازن است و برای فهمیدن اینکه آیا روش LS-TSVM قادر به تشخیص هر دو کلاس می باشد، از معیار های مذکور استفاده شده است.

$$Accuracy = \frac{(TP + TN)}{N} \quad (8)$$

$$Recall(R) = \frac{TP}{(TP + FN)} \quad (9)$$

$$Precision(P) = \frac{TP}{(TP + FP)} \quad (10)$$

<sup>1</sup> Perpendicular distance

<sup>2</sup> Kernel function

<sup>3</sup> Accuracy

<sup>4</sup> Recall

<sup>5</sup> Precision

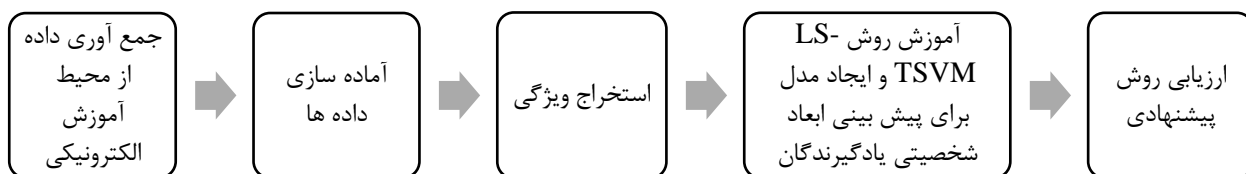


$$F1 = \frac{2 \times PR}{(P + R)} \quad (11)$$

متغیر TP بیانگر تعداد نمونه‌های کلاس 1 است که به درستی کلاس 1 تشخیص داده شده، متغیر TN نیز تعداد نمونه‌های کلاس 1- است که به درستی کلاس 1- تشخیص داده شده، متغیر FP، تعداد نمونه‌های کلاس 1- است که به اشتباه کلاس 1 تشخیص داده شده و همچنین متغیر FN تعداد نمونه‌های کلاس 1 است که به اشتباه کلاس 1- تشخیص داده شده است. متغیر N نیز برابر با تعداد کل نمونه‌های مجموعه داده می‌باشد.

#### ۴-۵ فرآیند کلی این پژوهش

در شکل ۷، فرآیند کلی انجام این پژوهش نشان داده شده است. جمع آوری و آماده‌سازی داده‌ها در بخش ۴ توضیح داده شد. بعد از آماده سازی داده‌ها برای آموزش دادن الگوریتم یادگیری نیاز به استخراج ویژگی که ۱۹ ویژگی استخراج شد که در بخش ۱-۵ توضیح داده شده است. با ویژگی‌های استخراج شده، روش LS-TSVM را آموزش می‌دهیم که یک مدل برای پیش‌بینی ابعاد شخصیتی یادگیرندگان ایجاد می‌شود. در نهایت کارایی مدل ایجاد شده با معیارهای ارزیابی ذکر شده در بخش ۳-۵ مورد سنجش قرار می‌گیرد. تحلیل و نتایج ارزیابی مدل ایجاد شده در بخش ۲-۶ آمده است.



#### شکل ۷ - فرآیند کلی انجام این پژوهش

#### ۶. نتایج ارزیابی مدل ایجاد شده برای پیش‌بینی ابعاد شخصیتی یادگیرندگان

در بخش ۶ ابتدا نحوه پیاده سازی روش پیشنهادی و اجرای الگوریتم‌ها توضیح داده شده است. سپس نتایج ارزیابی روش پیشنهادی، تحلیل آن و مقایسه آن با سایر الگوریتم‌ها آورده شده است.

#### ۶-۱ نحوه پیاده‌سازی روش پیشنهادی و اجرای الگوریتم‌ها

برای آماده سازی داده‌ها و پیاده سازی روش یادگیری ماشین بردار پشتیبان دو قلو کمترین مربعات از زبان برنامه‌نویسی Python<sup>۱</sup> نسخه ۳/۶ و همچنین از کتابخانه‌های Numpy و Scikit-learn استفاده شده است. برای پیش‌بینی ابعاد شخصیتی دانشجویان از نسخه غیر خطی LS-TSVM با تابع هسته گوسی استفاده شده است که پیدا کردن

<sup>1</sup> <http://www.python.org>

بهترین پارامترهای آن با جستجوی شبکه‌ای<sup>۱</sup> انجام شده است. پارامترهای C1 و C2 بین بازه  $2^{-7}$  تا  $2^2$  و پارامتر تابع گوسی نیز بین  $2^{-20}$  تا  $2^4$  است. روش پیشنهادی این پژوهش با ماشین بردار پشتیبان مقایسه شده است. برای اجرای SVM از نرم افزار Weka نسخه ۳/۸/۱ استفاده شده است [۱۲]. همچنین تابع هسته چند جمله‌ای برای SVM استفاده شده است.

## 6-2 نتایج ارزیابی روش پیشنهادی

در جدول ۲ نتایج ارزیابی روش پیشنهادی این پژوهش و ماشین بردار پشتیبان آمده است. هر دو روش با ویژگی‌های بیان شده در جدول ۱ و روی ۴ طبقه بندی دودویی شامل فکری - احساسی، برونگرا - درونگرا، قضاوتی - ادراکی و حسی - شهودی آزمایش شده است. همچنین هر دو روش با اعتبار سنج ضربدری<sup>۲</sup> به صورت 2-fold ارزیابی شده است.

**جدول ۲- نتایج ارزیابی بر روی ۴ طبقه بندی دودویی. واحد اعداد درصد می‌باشد.**

میانگین دو کلاس F1	کلاس ۱- F1	کلاس ۱- صحت	کلاس ۱- بازخوانی	کلاس ۱ F1	کلاس ۱ صحت	کلاس ۱ بازخوانی	دقت	نوع طبقه بندی کلاس ۱، کلاس ۱-
ماشین بردار پشتیبان با تابع هسته چند جمله‌ای								
۸۸/۱	۸۳/۳	۸۳/۳	۸۳/۳	۹۲/۹	۹۲/۹	۹۲/۹	۹۰	فکری، احساسی
۸۹/۶۵	۹۵/۱	۹۰/۶	۱۰۰	۸۴/۲	۱۰۰	۷۲/۷	۹۲/۵	برونگرا، درونگرا
۸۳/۵۵	۷۸/۶	۸۴/۶	۷۳/۳	۸۸/۵	۸۵/۲	۹۲	۸۵	قضاوتی، ادراکی
۸۴/۴	۷۵	۱۰۰	۶۰	۹۳/۸	۸۸/۲	۱۰۰	۹۰	حسی، شهودی
۸۶/۴	۸۳	۸۹/۶	۷۹/۱	۸۹/۸	۹۱/۵	۸۹/۴	۸۹/۳	میانگین
روش پیشنهادی این مقاله								
<b>97/15</b>	96/15	92/86	100	98/15	100	96/43	<b>97/5</b>	فکری، احساسی
<b>96/86</b>	98/28	100	96/67	95/45	91/67	100	<b>97/5</b>	برونگرا، درونگرا
<b>96/86</b>	95/45	91/67	100	98/28	100	96/67	<b>97/5</b>	قضاوتی، ادراکی
<b>89/24</b>	83/33	87/5	80	95/16	93/75	96/67	<b>92/5</b>	حسی، شهودی
<b>95</b>	<b>93/3</b>	<b>93</b>	<b>94/1</b>	<b>96/7</b>	<b>96/3</b>	<b>97/4</b>	<b>96/2</b>	میانگین

ماشین بردار پشتیبان روی ۴ طبقه بندی با در نظر گرفتن میانگین F1 دو کلاس، عملکرد رضایت بخشی دارد. با این حال روی طبقه بندی حسی - شهودی، مقدار بازخوانی بعد شهودی برای SVM برابر با ۶۰ درصد می‌باشد که نسبت به LS-TSVM، ۲۰ درصد ضعیف‌تر است. زیرا ماشین بردار پشتیبان دو قلو برای مجموعه داده‌های نامتوازن مناسب‌تر است [۱۱]. با وجود اینکه دقت ماشین بردار پشتیبان روی ۴ طبقه بندی بالاتر از ۸۵ درصد بوده است، با در نظر گرفتن

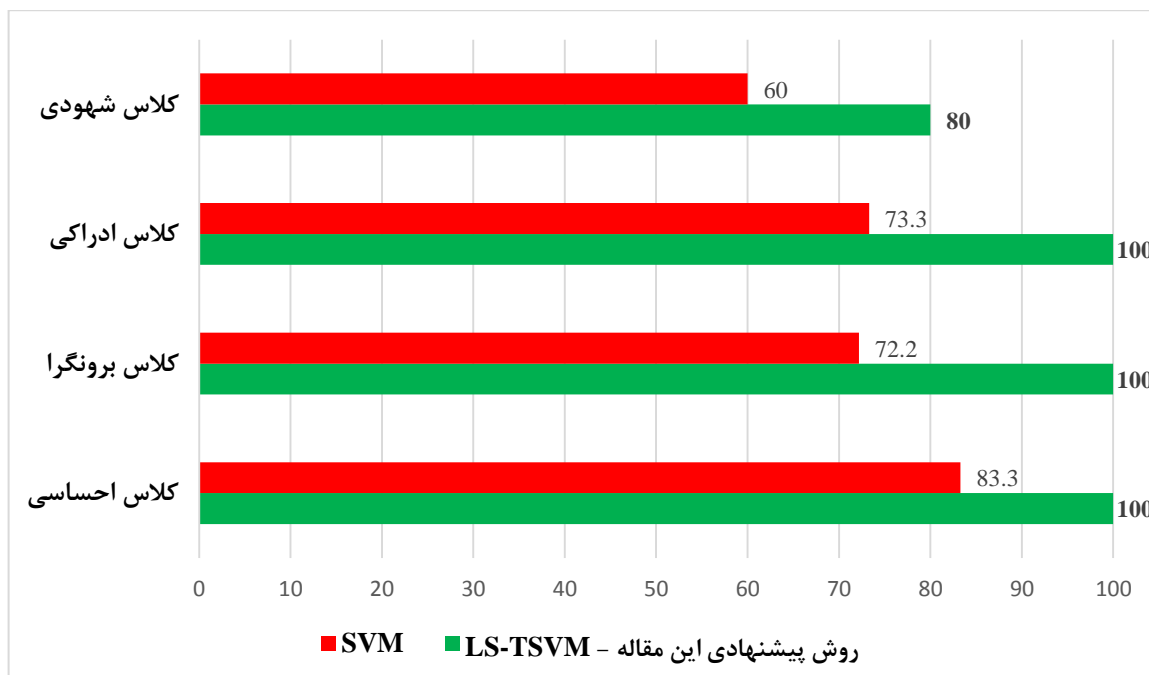
<sup>1</sup> Grid search

<sup>2</sup> Cross-validation

سایر معیارهای ارزیابی به خصوص معیار بازخوانی، عملکرد این روش نشان می‌دهد که قابلیت تبدیل به یک مدل قابل اطمینان را ندارد.

روش پیشنهادی این مقاله، در هر ۴ طبقه بندی دقت بالای ۹۰ درصد بدست آورده است. همچنین میانگین F1 دو کلاس روش پیشنهادی نیز در هر ۴ طبقه بندی به جز یک مورد بالای ۹۰ درصد می‌باشد. نتایج روش پیشنهادی این مقاله نشان می‌دهد که می‌توان از این روش یک مدل برای پیش‌بینی ابعاد شخصیتی دانشجویان جهت طراحی شیوه تدریس و آموزش بهتر، ساخت. روش LS-TSVM در طبقه بندی حسی - شهودی نسبت به سایر طبقه‌بندی‌ها ضعیف-تر عمل کرده و میانگین F1 دو کلاس برابر با ۸۹/۲۴ درصد شده است. به دلیل اینکه مجموعه داده حسی - شهودی همانطور که در بخش ۲-۴ نشان داده شد، بسیار نامتوازن است و فقط ۲۵ درصد نمونه‌ها متعلق به کلاس شهودی است که همین موضوع باعث شده است که مقدار بازخوانی برای کلاس شهودی به ۸۰ درصد برسد. با این حال روش LS-TSVM با وجود نامتوازن بودن داده‌ها، عملکرد خوبی در هر ۴ طبقه بندی دارد.

برای مقایسه بهتر روش پیشنهادی این پژوهش با ماشین بردار پشتیبان، مقدار بازخوانی کلاس‌های احساسی، برونگرا، ادراکی و شهودی برای هر دو روش در نظر گرفته می‌شود. توزیع کلاس‌های ذکر شده در بخش ۲-۴ آورده شده است. کلاس احساسی، برونگرا، ادراکی و شهودی به ترتیب ۳۰ درصد، ۲۷ درصد، ۳۸ درصد و ۲۵ درصد نمونه‌ها را در مجموعه داده خود شامل می‌شوند. در شکل ۸ مقدار بازخوانی کلاس‌های ذکر شده برای روش پیشنهادی این مقاله و ماشین بردار پشتیبان آورده شده است.



شکل ۸ - مقایسه روش پیشنهادی این مقاله با SVM بر اساس معیار بازخوانی

در این مقایسه، ماشین بردار پشتیبان فقط در کلاس احساسی عملکرد مناسب و خوبی دارد و روی سایر کلاس‌ها به دلیل نامتوازن بودن داده‌ها عملکرد مناسبی ندارد. در کلاس شهودی، ماشین بردار پشتیبان فقط ۶۰ درصد یادگیرندگان با بعد شهودی را به درستی تشخیص داده است. در مقابل روش LS-TSVM تمام نمونه‌های کلاس احساسی، برونگرا و

ادراکی را به درستی تشخیص داده است. با وجود اینکه مجموعه داده حسی - شهودی بسیار نامتوازن است، روش LS-TSVM روی کلاس شهودی نیز عملکرد مناسبی دارد.

روش LS-TSVM نسبت به SVM روی کلاس های احساسی، برونگرا، ادراکی و شهودی بر اساس معیار بازخوانی به ترتیب ۱۶/۷ درصد، ۲۷/۸ درصد، ۲۶/۷ درصد و ۲۰ درصد بهتر عمل کرده است که قابل توجه می باشد. در مجموع مقایسه انجام گرفته در شکل ۸ نشان می دهد که روش LS-TSVM نسبت SVM روی مجموعه داده های نامتوازن عملکرد بهتری دارد و انتخاب روش LS-TSVM در این پژوهش به عنوان طبقه بندی کننده را توجیه می کند.

#### ۷. نتیجه گیری و کارهای آینده

پیش بینی ابعاد شخصیتی یادگیرندگان در یک محیط آموزشی به منظور بهبود شیوه تدریس، آموزش و افزایش کارایی دانشجویان اهمیت دارد. در این مقاله برای تحقق این امر، از یک محیط آموزش الکترونیکی، داده هایی جمع آوری شده است. بعد از آماده سازی داده ها، ۱۹ ویژگی برای پیش بینی ابعاد شخصیتی یادگیرندگان استخراج شده است. به دلیل نامتوازن بودن مجموعه داده این پژوهش از ماشین بردار پشتیبان دو قلو کمترین مربعات به عنوان طبقه بندی کننده استفاده شده است. نتایج نشان می دهد که روش پیشنهادی این پژوهش عملکرد خوبی دارد و از ماشین بردار پشتیبان بهتر عمل کرده است. همچنین نتایج بدست آمده این اطمینان را می دهد که می توان از روش پیشنهادی این پژوهش یک مدل برای پیش بینی ابعاد شخصیتی یادگیرندگان در یک محیط آموزش الکترونیکی ایجاد کرد. کارهایی که برای پژوهش آینده می توان انجام داد، عبارتند از:

۱. داده های بیشتری از یک محیط آموزشی جدید جمع آوری شود و کارایی مدل پیشنهادی این پژوهش با داده های جدید سنجیده شود.
۲. در این پژوهش، در یک محیط آموزش الکترونیکی تیپ شخصیتی افراد پیش بینی می شود. در پژوهش های آینده می توان دامنه مسئله پیش بینی تیپ شخصیتی افراد را تغییر داد. می توان در بازی های کامپیوتری و یا فروشگاه های اینترنتی، تیپ شخصیتی بازیکنان و یا خریداران را پیش بینی کرد.

#### ۸. مراجع

1. Fatahi, S. and H. Moradi, *A fuzzy cognitive map model to calculate a user's desirability based on personality in e-learning environments*. Computers in Human Behavior, ۲۰۱۶. ۶۳: p. ۲۸۱-۲۷۲
2. Ma, A. and G. Liu, *Neural Networks in Predicting Myers Brigg Personality Type From Writing Style*. ۲۰۱۷
3. Barrick, M.R. and M.K. Mount, *The big five personality dimensions and job performance: a meta-analysis*. Personnel psychology, ۱۹۹۱. ۴۴(۱): p. ۲۶-۱
4. Felder, R.M., G.N. Felder, and E.J. Dietz, *The effects of personality type on engineering student performance and attitudes*. Journal of engineering education, ۲۰۰۲. ۹۱(۱): p. ۱۷-۳
5. Gong, W. and W. Wang. *Application research of support vector machine in E-Learning for personality*. in *Cloud Computing and Intelligence Systems (CCIS)*, ۲۰۱۱ IEEE International Conference on. ۲۰۱۱. IEEE.

6. Kalghatgi, M.P., M. Ramannavar, and N.S. Sidnal, *A neural network approach to personality prediction based on the big-five model*. International Journal of Innovative Research in Advanced Engineering (IJIRAE), ۲۰۱۵. ۲(۸): p. ۵۶-۶۳.
7. Fatahi, S., H. Moradi, and E. Farmad, *Behavioral Feature Extraction to Determine Learning Styles in e-Learning Environments*. International Association for Development of the Information Society, ۲۰۱۵.
8. Kumar, M.A. and M. Gopal, *Least squares twin support vector machines for pattern classification*. Expert Systems with Applications, ۲۰۰۹. ۳۶(۴): p. ۷۵۳۵-۷۵۴۳.
9. Fatahi, S., H. Moradi, and L. Kashani-Vahid, *A survey of personality and learning styles models applied in virtual environments with emphasis on e-learning environments*. Artificial Intelligence Review, ۲۰۱۶. ۴۶(۳): p. ۴۱۳-۴۲۹.
10. Cortes, C. and V. Vapnik, *Support-vector networks*. Machine learning, ۱۹۹۵. ۲۰(۳): p. ۲۷۳-۲۹۷.
11. Khemchandani, R. and S. Chandra, *Twin support vector machines for pattern classification*. IEEE Transactions on pattern analysis and machine intelligence, ۲۰۰۷. ۲۹(۵).
12. Witten, I.H., et al., *Data Mining: Practical machine learning tools and techniques*. ۲۰۱۶: Morgan Kaufmann.