



امکان سنجی قابلیت طیف سنجی مرئی/فروسرخ نزدیک (Vis/NIR) در طبقه بندی نمونه های لیموترش طی دوره انبارمانی با روش های شناسایی LDA, PCA و SVM

نیلوفر گودرزی^۱، سارا موحد^{۲*}، محمد جواد شکوری^۳، حسین احمدی چناربن^۴

- ۱- دانشجوی دکتری، گروه علوم و صنایع غذایی، دانشکده کشاورزی، واحد تهران شمال، دانشگاه آزاد اسلامی، تهران، ایران.
- ۲- دانشیار، گروه علوم و صنایع غذایی، دانشکده کشاورزی، واحد ورامین- پیشوا، دانشگاه آزاد اسلامی، ورامین، ایران.
- ۳- استادیار، گروه علوم و صنایع غذایی، دانشکده کشاورزی، واحد تهران شمال، دانشگاه آزاد اسلامی، تهران، ایران.
- ۴- استادیار، گروه زراعت و اصلاح نباتات، دانشکده کشاورزی، واحد ورامین- پیشوا، دانشگاه آزاد اسلامی، ورامین، ایران.

چکیده	اطلاعات مقاله
<p>امروزه روند افزایشی ضایعات مواد غذایی و محصولات کشاورزی یکی از چالش‌های جدی اکثر کشورها، به ویژه کشورهای در حال توسعه محسوب می‌شود لذا یکی از سیاست‌های جدی دولت‌ها در امر امنیت مواد غذایی، کاهش ضایعات و حفظ کیفیت محصولات کشاورزی است. تاکنون از روش‌های متعددی برای سنجش کیفیت محصولات کشاورزی استفاده شده است که تنها برخی از آنها از لحاظ فنی و صنعتی توجیه پذیرند. روش طیف سنجی مرئی/مادون قرمز نزدیک (Vis/NIR) از جمله روش‌هایی است که به دلیل سرعت و دقت بالا در ارزیابی خصوصیات کیفی محصولات کشاورزی مورد توجه و استفاده قرار گرفته است. در این راستا، در پژوهش حاضر از طیف سنجی مرئی/فروسرخ نزدیک به منظور سنجش تغییرات کیفی و طبقه‌بندی نمونه‌های لیموترش وارسته کی لایم، طی دوره انبارمانی (۱۰، ۲۰ و ۳۰ روز) استفاده گردید. در ادامه به منظور تحلیل ویژگی‌های کیفی و طبقه بندی داده‌های مستخرج از NIR، از روش های شناسایی الگو شامل تحلیل مؤلفه های اصلی (PCA)، تحلیل تفکیک خطی (LDA) و ماشین بردار پشتیبان (SVM) استفاده شد. نتایج بدست آمده نشان داد که طیف‌سنجی مرئی/فروسرخ نزدیک (Vis/NIR) قادر به تفکیک نمونه‌های لیموترش بر اساس مدت زمان نگهداری در انبار است. هرچند روش‌های LDA، PCA و SVM توانستند با دقت خوبی نمونه‌های لیموترش را با توجه به ویژگی‌های کیفی دسته‌بندی کنند، اما روش‌های LDA و SVM با دقت ۱۰۰٪ از دقت و برازش مطلوب‌تری برخوردار بودند. همچنین، طبق نتایج، تابع درجه ۲، به عنوان بهترین تابع برای ساخت مدل‌های دسته‌بندی به روش‌های LDA و SVM تعیین و معرفی گردیده است.</p>	<p>تاریخ های مقاله :</p> <p>تاریخ دریافت: ۱۴۰۰/۰۵/۱۴</p> <p>تاریخ پذیرش: ۱۴۰۰/۰۷/۱۰</p> <p>کلمات کلیدی:</p> <p>فروسرخ نزدیک، طیف سنجی مرئی، لیموترش، دوره انبارمانی.</p> <p>DOI: 10.52547/fsct.18.120.26</p> <p>DOR: 20.1001.1.20088787.1400.18.120.27.9</p> <p>* مسئول مکاتبات: movahhed@iauvaramin.ac.ir</p>

۱- مقدمه

مرکبات در بیش از ۵۰ کشور جهان بصورت تجاری کشت می‌گردد و کالای اصلی در تجارت بین‌الملل به حساب می‌آید [۱،۲]. مرکبات جزء محصولات نافرزاگرا بوده و تنها در مرحله رسیدن کامل جهت مصارف خوراکی برداشت می‌شوند. این مرحله ملزم به برآوردن انتظارات مصرف‌کنندگان می‌باشد، این انتظارات نه تنها شامل بهترین ظاهر، طعم و خواص تغذیه‌ای است بلکه ایمنی و پایداری هر محصول را نیز شامل می‌شود [۳]. اصلی‌ترین ویژگی در نظر گرفته شده برای سنجش کیفیت خارجی مرکبات بطورکلی ظاهر، اندازه، وزن و رنگ می‌باشد. در این میان سنجش ویژگی‌هایی همچون محتوای جامد محلول (SSC)، اسیدیته قابل تیتر (TA)، شاخص رسیدگی (MI) و عدم وجود نقایص داخلی از مهمترین مواردی است که جهت تعیین کیفیت داخلی میوه جاتی چون لیموترش مورد بررسی قرار می‌گیرد. اگرچه میزان سفتی میوه از نظر کمی بیان نشده است، اما به عنوان یکی از مهمترین پارامترهای سنجش کیفیت داخلی میوه جات به حساب می‌آید زیرا یکی از فاکتورهای مهم در عملیات پس از برداشت، حمل و نقل و انبارمانی می‌باشد [۴]. در میان مرکبات، لیموترش یکی از مهمترین و اقتصادی‌ترین محصولات باغی در جنوب ایران است که در سطح وسیعی معادل ۴۱۸۰۰ هکتار، کشت می‌گردد و استان هرمزگان با ۷۰٪ کل تولید، قطب لیموترش در ایران محسوب می‌شود. بر اساس آمار سازمان خواربار و کشاورزی ملل متحد (FAO)، ایران با تولید سالانه ۶۱۵۰۰۰ تن، در میان ۱۰ کشور برتر تولیدکننده لیموترش در جهان قرار دارد [۵]. لایم با نام علمی *Citrus aurantifolia* L. از گونه‌های اولیه مرکبات بوده که دارای دو نوع میوه‌ی اسیدی و غیر اسیدی می‌باشد. از میوه‌های غیر اسیدی می‌توان به لیمو شیرین اشاره نمود اما ارقام اسیدی شامل میوه‌های بزرگ (Persian lime) و کوچک (Mexican lime and Key lime) می‌باشند. در ایران رقم Mexican lime به نام لیمو جهرمی شناخته می‌شود [۶]. اجزاء زیست فعال لیمو از قبیل فلاونونوئیدها، ویتامین‌ها، موادمعدنی، فیبر و روغن‌های اسانسی بوده که از آنها در صنایع غذایی، آرایشی و داروسازی استفاده می‌کنند. پوست لیمو که ۲۵-۳۰ درصد وزن کل میوه را شامل می‌شود، ماده اصلی محصولات فرعی را تشکیل می‌دهد. همچنین وجود مواد

بیواکتیو (فلاونونوئیدها و ویتامین C) در قسمت آلبیدوی پوست لیمو باعث تبدیل شدن این بخش به یک منبع فیبری شده است. لیموترش منبع غنی از کاروتنوئیدها می‌باشد. از سوی دیگر روغن‌های اسانسی لیموترش به عنوان طعم دهنده و معطرکننده در صنایع غذایی و داروسازی به کار می‌روند که ناشی از خصوصیات ضد میکروبی و ضد قارچی آنهاست. آب لیموترش یک منبع مهم اسیدسیتریک می‌باشد که به‌طور مؤثری می‌تواند برای از بین بردن یون‌های فلزات سنگین در محلول‌های آبی استفاده شود [۷]. بطور کلی نیاز به ارتقاء مدیریت و پایداری زنجیره تأمین مرکبات با فناوری‌های هوشمند و غیر مخرب، روز به روز احساس می‌شود تا ارزیابی سریع، عینی و دقیقی از کیفیت مرکبات و رسیدن آنها بر روی درخت و حتی بعد از برداشت صورت گیرد و جایگزین روش‌های متداول و رایج گردد. در واقع هدف اصلی، ارائه بهترین محصولات به بازار و کمک به کاهش جریان ضایعات مواد غذایی از جمله میوه جات و سبزیجات در سراسر جهان است [۸، ۹]. با توجه به اینکه چه مقدار از جمعیت جهان فاقد غذای ایمن هستند که ضروریات پر اهمیتی از جمله مواد مغذی و ویتامین‌ها را دارا باشد و از سوء تغذیه جلوگیری کند، بیش از پیش اهمیت فناوری‌های جدید را نمایان می‌کند. تعداد زیادی از گزارشات منتشر شده در دو دهه گذشته نشان دهنده تحقیقات بسیار کارآمد و پویا بر روی روش‌های غیر مخرب جهت ارزیابی کیفیت و پارامترهای رسیدن انواع مختلفی از میوه جات از جمله مرکبات بوده است. در میان آنها، روش طیفسنجی مرئی/مادون قرمز نزدیک (Vis/NIR) یکی از مناسب‌ترین و پیشرفته‌ترین روش‌های غیر مخرب نام برده شده است. که در حال حاضر برای نظارت و کنترل چندین محصول باغی مورد استفاده قرار می‌گیرد [۱۰، ۱۱] و جزء روش‌هایی است که با توجه به سرعت و دقت بالا به منظور ارزیابی خصوصیات کیفی محصولات کشاورزی مورد توجه قرار گرفته است [۱۲-۱۷]. طیفسنجی مادون قرمز نزدیک، ابزار قدرتمندی است که می‌تواند عرضه محصولات تازه را برجسته تر نماید و ضایعات مواد غذایی را کاهش دهد. علاوه بر این، می‌تواند رقابت پذیری را در حین تثبیت قیمت افزایش دهد. طیفسنجی NIR در حال حاضر در دستگاه‌های قابل حمل موجود می‌باشد که می‌تواند کیفیت محصولات تازه را کنترل نماید. بطوریکه در برنامه‌های

فضای آزمایشگاهی دارد اما طیف سنجی NIR اکنون مینیاتوری شده و در ابزارهای کوچک و قابل حمل در دسترس است و نه تنها کاربر پسند است بلکه به اندازه کافی ساده و راحت می باشد. میزان ماده خشک در هنگام برداشت، ابزار خوبی جهت پیش بینی شاخص رسیدن میوه بعد از برداشت است. با این حال مراحل زیادی بین برداشت تا فروش وجود دارد که می تواند بر کیفیت محصولات تازه نیز تأثیر بگذارد نظیر فرآیند شستشو، درجه بندی، حمل و نقل و غیره. میوه جات و سبزیجات زنده هستند و ویژگی های آنها در طول این فرآیندها تغییر می کند، بنابراین طیف سنجی NIR به عنوان یک معیار کنترل کیفیت در همه این مراحل مفید است. به عنوان مثال میزان رسیدگی خرما توسط بریکس، تردی، سختی و رنگ زرد آن مشخص می گردد. طیف سنجی NIR می تواند تمام مراحل فوق الذکر رسیدن خرما را با استفاده از میزان مواد قندی محلول و میزان رطوبت مشخص نماید [۳۶].

پارامترهای رسیدن میوه جات چندین بار در زنجیره تأمین جهت درجه بندی میوه ها در مراحل مختلف بررسی می شوند. میوه جات فرازگرا زمانیکه کاملاً رسیده نیستند، برداشت می شوند اما محصولات نافرزگرا بعد از رسیدن کامل برداشت می شوند. ابزارهای مبتنی بر طیف سنجی NIR برای درجه بندی محصولات تازه با توجه به شاخص های رسیدن آنها مفید می باشد. به عنوان مثال زمانیکه گوجه فرنگی ها برداشت می شوند و عملیات درجه بندی روی آنها صورت نمی گیرد منجر به ضایعات ۱۰ درصد از محصولات برداشت شده می شود. زیرا گوجه فرنگی های رسیده یا پوسیده باعث تولید اتیلن می شوند که سبب رسیدن بیشتر مابقی محصولات می گردند [۳۷]. طیف سنجی NIR می تواند به طرق مختلف برای تشخیص تغلب در بازار مورد استفاده قرار گیرد از جمله: ۱/ یافتن منشأ جغرافیایی محصولات ۲/ روش های کشاورزی بکار گرفته شده جهت تشخیص محصولات معمولی از محصولات ارگانیک ۳/ شناسایی محصولات تحت نظارت نظیر برخی گونه های سیب در ایتالیا که جهت بهبود رنگ و طعم باید فرآیند قرمز شدن را طی کنند [۳۸]. شرایط انبارمانی در طی دمای پایین، حضور گازهای کنترل شده و همچنین مدت زمان نگهداری بر کیفیت محصول نهایی تأثیرگذار است. از سوی دیگر، شرایط نگهداری می تواند بر مدت زمان ذخیره سازی میوه تأثیرگذار باشد. زمانیکه میوه ها در دمای پایین

کاربردی اعم از درجه بندی خودکار خطی اجراء می گردد و ارزیابی بیش از ۱۰ تا ۱۲ میوه در هر ثانیه را انجام می دهد، همچنین در نور کامل خورشید و دماهای مختلف محیط قابل استفاده می باشد. این فناوری همچنین برای کنترل تغییرات کیفیت تولید محصولات در طول زنجیره تأمین مفید می باشد. علاوه بر کشاورزان، بسیاری از ذینفعان در زنجیره تأمین می توانند از فناوری استفاده کنند نظیر توزیع کنندگان، خرده فروشان و مصرف کنندگان [۱۸-۲۰]. اشعه مادون قرمز نزدیک (NIR)، اولین بار توسط Friedrich Wilhelm Herschel در سال ۱۸۰۰ میلادی کشف شد، و اولین کاربرد آن در صنعت کشاورزی جهت اندازه گیری رطوبت دانه در اواخر دهه ۱۹۶۰ میلادی بود [۲۱].

از مزیت های عمده طیف سنجی فرسوخ نزدیک می توان به موارد زیر اشاره کرد:

- عدم نیاز به آماده سازی نمونه
- عدم تولید پسماند
- اندازه گیری سریع
- امکان آنالیز نمونه در داخل ظروف شیشه ای
- قابلیت به کارگیری از راه دور
- هزینه آنالیز
- قابلیت به کارگیری در مزرعه
- عدم نیاز به مهارت های خاص [۲۲].

بطورکلی، تکنولوژی های روش غیر مخرب شامل بینی الکترونیکی [۲۳، ۲۴، ۲۵]، طیف سنجی مادون قرمز نزدیک (Vis/NIR) [۲۶، ۲۷، ۲۸]، تکنیک های اولتراسونیک [۲۹، ۳۰]، تکنیک ماشین بینایی [۳۱، ۳۲، ۳۳] و ادغام چند روش غیر مخرب [۳۴، ۳۵] می باشند، اما در این بین، تکنیک طیف سنجی مادون قرمز نزدیک با سرعت و دقتی بالا، هزینه پایین اجراء، رویکرد امیدوارانه ای را برای تشخیص کیفیت محصولات کشاورزی ارائه می دهد. یکی از الزامات کنترل کیفیت میوه جات و سبزیجات آن است که اندازه گیری ها باید غیر مخرب باشند تا منجر به حفظ محصولات و کاهش ضایعات گردند. از جمله روش های متداولی همچون آون، میکروویو جهت سنجش ماده خشک و یا رفراکتومتر جهت سنجش میزان شیرینی در محصولات، اگرچه غیر مخرب هستند اما زمان بر بوده و برای انجام تجزیه و تحلیل شیمیایی نیاز به آموزش دارند لذا آماده سازی نمونه در آنها نیاز به

محصولات باغی همچون غده‌های نظیر سیب زمینی و در حد کمتر پیاز و قارچ نیز مورد هدف طیفسنجی NIR قرار می‌گیرند اما استفاده از آن در سبزیجات برگی تا به امروز بسیار کمتر بوده است [۴۰]. اولین کاربردهای تکنولوژی طیفسنجی NIR پس از برداشت، ارزیابی رسیده بودن و تشخیص عیوب داخلی محصولات می‌باشد که این امر با تشخیص غیر مستقیم محتوای قند در میوه‌ها توسط تعیین محتوای مواد جامد محلول و ماده خشک امکان‌پذیر است. سبب جزء اولین میوه‌هایی بوده است که کاربرد NIR، پس از برداشت روی آن مورد تحقیق و بررسی قرار گرفته است [۴۱]. در میوه‌های نافرازگرا مانند مرکبات، انگور و غیره میزان قند و اسید قبل و بعد از برداشت نسبتاً پایدار است. نسبت بین دو پارامتر نشان داده شده توسط اسیدیته کل و محتوای ماده جامد، کلید مقبولیت مصرف‌کننده و بطور گسترده به عنوان معیار بلوغ در این میوه‌جات می‌باشد. طیف‌های NIR حاصل تابش و برهم‌کنش نمونه هستند و بازتاب آنها تعیین‌کننده خواص فیزیکی و شیمیایی است. ویژگی‌هایی مانند سفتی، آبدار بودن، وزن و رنگ از جمله ویژگی‌های فیزیکی است [۴۲].

مطالعات گذشته بیانگر اهمیت بالای روش طیفسنجی به عنوان یک روش سریع، ساده، کم‌هزینه با دقت قابل قبول در سنجش کیفیت محصولات کشاورزی است. به همین دلیل، استفاده از آن بصورت‌های مختلف و برای محصولات گوناگون روز به روز در حال گسترش است. اما تاکنون پژوهشی که در آن به دسته‌بندی لیموترش مبتنی بر روش طیفسنجی و با استفاده از روش‌های طبقه‌بندی پرداخته شده باشد، یافت نشد. بنابراین، در این پژوهش ماندگاری لیموترش‌ها و طبقه‌بندی آنها با استفاده از طیفسنجی مرئی/فروسرخ نزدیک (Vis/NIR) بدون سنجش و اندازه‌گیری خواص کیفی با استفاده از روش‌های شناسایی الگو نظیر PCA، LDA و SVM انجام پذیرفت. در واقع تجزیه و تحلیل داده‌های چند متغیرها از جمله روش‌های آماری است که بر مشاهده همزمان، بهره‌برداری و پردازش چندین متغیر آماری به منظور استخراج اطلاعات سیستمیک مربوطه استوار می‌باشد. از این رو اهداف خاص این مطالعه عبارتند از: ۱/ مقایسه عملکرد روش‌های پیش‌پردازش طیفی مختلف ۲/ آنالیز توزیع متغیرهای انتخاب شده به وسیله الگوریتم‌های مختلف انتخاب متغیر ۳/ مقایسه دقت پیش‌بینی مدل ساخته شده به

نگهداری می‌شوند، با کاهش تولید اتیلن مدت زمان نگهداری آنها نیز افزایش می‌یابد، با این حال باید مراقب بود که دمای هوا خیلی پایین نیاید زیرا باعث ایجاد صدمات یخ‌زدگی در میوه‌جات می‌شود. با استفاده از طیفسنجی NIR می‌توان دریافت که گونه و تنوع محصول چه میزان در شرایط اتمسفری مختلف بدون تأثیر بر کیفیت می‌تواند ذخیره گردد. به عنوان مثال خرما می‌تواند به ترتیب ۲۰، ۴۰ و ۱۲۰ روز در شرایط دمای اتاق، سردخانه و اتمسفر کنترل شده نگهداری شود [۲۰]. اهمیت تغذیه‌ای میوه‌جات و سبزیجات، اطمینان از پذیرش آنها توسط مصرف‌کنندگان از اهداف اصلی دوران بعد از برداشت می‌باشد که سبب تشویق و توسعه فناوری برای نظارت بر ویژگی‌های کیفی است. طیفسنجی مادون قرمز نزدیک پتانسیل بالایی برای دستیابی به این اهداف دارد. تکنیک‌های غیر مخرب برای تجزیه و تحلیل کیفیت میوه ممکن است در مراحل مختلف زنجیره تولید جالب باشد. اولین مورد مربوط به عملیات قبل از برداشت است. توانایی نظارت یا تجزیه و تحلیل فوری پارامترهای کیفیت میوه در مزرعه، یکی از مزایای بزرگ توسعه استراتژی‌های برداشت می‌باشد. از انتقال نمونه به آزمایشگاه جهت انجام تست‌های آزمایشگاهی اجتناب می‌گردد در نتیجه در زمان و هزینه صرفه‌جویی به عمل می‌آید. مرحله بعدی رسیدگی و بسته‌بندی میوه‌جات می‌باشد، جاییکه درجه بندی خودکار میوه با توجه به برخی پارامترهای کیفیت انجام می‌گردد. مرحله دیگر، زنجیره توزیع مواد غذایی می‌باشد. در این فاز توزیع‌کنندگان و مصرف‌کنندگان علاقمند به تأیید راحت و سریع محصول به دست آمده و مطابقت آن با ویژگی‌های مورد نظر هستند، این گزینه می‌تواند شفافیت در بازاریابی و رقابت پذیری در قیمت‌فراپندها را به همراه داشته باشد. با توجه به دلایل ذکر شده، در حال حاضر علاقه زیادی به تجزیه و تحلیل غیر مخرب کیفیت میوه‌جات وجود دارد [۳۹]. پتانسیل طیفسنجی مادون قرمز نزدیک، سریع، غیر مخرب و تکنیکی چند پارامتری است که سبب می‌شود بیشتر کاربردهای آن بعد از برداشت در محصول نمایان شود. علاوه بر این، میوه‌جاتی از جمله محصولات باغی مانند گوجه‌فرنگی، خربزه، هندوانه، لیموترش، لعل و سایر موارد متداولترین هدف تجزیه و تحلیل طیفسنجی مادون قرمز، پس از برداشت به دلیل مشخص بودن تغییرات در ویژگی‌های تکاملی آنها می‌باشند. دیگر

منبع نور به نحوی قرار داده شد که زاویه بین آشکارساز و منبع نور ۴۵ درجه باشد.

۲-۲- آماده سازی و پیش پردازش طیف ها

داده های طیفی به دست آمده از اسپکترومتر، افزون بر اطلاعات نمونه شامل اطلاعات ناخواسته پس زمینه و نویزها نیز هستند. از این رو، برای دستیابی به مدل های واسنجی پایدار، دقیق و قابل اعتماد، نیاز به پیش پردازش داده های طیفی پیش از تدوین مدل دسته بندی متناسب با داده می باشد. ساخت مدل دسته بندی برای توصیف رابطه بین داده های طیفی و آزمایشگاهی یکی از مهمترین و سخت ترین مراحل طیف سنجی به شمار می رود و فرآیندی زمانبر است که معمولا با آزمون و خطا انجام می گیرد. به منظور اجرای مراحل پیش پردازش داده های طیفی نیاز به تبدیل داده های مد بازتابی به مد جذبی است که بدین منظور از رابطه ۱ استفاده شد که به قانون بیر-لمبرت مشهور است [۴۳].

$$A = \log_{10} / Tr_{\lambda} \quad (1)$$

که در آن: A: میزان جذب، Tr_{λ} : بازتاب یا عبور در طول موج λ

پس از این که طیف های بازتابی به دست آمده از همه نمونه ها به منظور ایجاد یک ارتباط خطی با غلظت مولکولی نمونه به طیف های جذبی تبدیل شدند، برای آماده سازی و پیش پردازش طیف ها از برخی روش های هموارسازی و کاهش نویز نظیر ساویتزکی-گولای (SG) و تصحیح خط مبنا (Baseline) استفاده شد.

۲-۳- تبیین مدل کالیراسیون

در آنالیز کمی و کیفی محصولات غذایی، ساخت و تدوین مدل واسنجی قابل اعتماد بسیار ضروری است. به کمک مدل می توان مشخصات نامعلوم نمونه ها را تعیین و پیش بینی نمود. در واقع هدف از تدوین مدل در این پژوهش ارزیابی ماندگاری لیموترش ها و طبقه بندی آنها می باشد. در این پژوهش از روش تحلیل تفکیک خطی^۱ (LDA)، تحلیل مؤلفه های اصلی^۲ (PCA) و ماشین بردار پشتیبان^۳ (SVM) استفاده گردید. ارزیابی مدل ساخته شده با روش اعتبارسنجی متقاطع کامل^۴ برای انتخاب مدلی با بیشترین توانایی دسته بندی، در بین

وسیله ترکیب پیش پردازش های طیفی مختلف و الگوریتم های انتخاب متغیر می باشد.

۲- مواد و روش ها

۲-۱- اندازه گیری با دستگاه طیف سنج فرورسرخ نزدیک (NIR)

تعداد ۱۲۰ نمونه لیموترش با وزن $20/50 \pm 2/50$ گرم از وارپته کی لایم، در مرحله ی رسیدگی کامل (زرد رنگ) به صورت تصادفی از بازار میوه تهران، در اوایل مرداد ماه سال ۹۹ خریداری شدند. سپس نمونه ها شماره گذاری گردیدند و در یخچالی با دمای ۳ درجه سلسیوس نگهداری شدند. طیف سنجی طی ۴ روز مختلف، صفر، ۱۰، ۲۰ و ۳۰ انجام پذیرفت. تعداد ۳۰ نمونه در هر روز به صورت تصادفی انتخاب شدند و قبل از طیف سنجی به مدت ۴ ساعت در محیط آزمایشگاه قرار گرفتند تا به دمای محیط برسند و اثرات منفی ناشی از اختلاف دما بر روی طیف خنثی گردد و سبب خطا نشود. به منظور طیف گیری از نمونه ها نیازی به هیچ گونه آماده سازی بر روی نمونه نمی باشد که این از مزایای طیف سنجی مرئی / فرورسرخ نزدیک است. طیف مرئی / فرورسرخ نزدیک با استفاده از فیبر دو مسیره که قسمت مرکزی برای انتقال نور از نمونه به آشکارساز و فیبر بیرونی برای انتقال نور از منبع تابش (لامپ هالوژن با توان ۵۰ وات) به نمونه می باشد، استفاده گردید، و در وضعیت انعکاسی و در ناحیه ۳۵۰ تا ۲۵۰۰ نانومتر با استفاده از اسپکترومتر مدل (StellarNet, USA) EPP2000NIR ساخت آمریکا و با تفکیک پذیری ۱ نانومتر توسط نرم افزار View Spec Pro ثبت گردید. در تمام طول آزمایش شرایط اتاق کاملا تاریک بود. از هر نمونه در سه جهت با هدف کاهش نویز حرارتی بر روی آشکارساز و افزایش دقت، تعداد ۳ اسکن متوالی گرفته شد و میانگین ۹ طیف حاصله، برای هر نمونه گزارش گردید. برای تهیه طیف بازتابی ابتدا دو طیف مرجع و تاریک به منظور تعیین بیشترین و کمترین حالت بازتاب برای سامانه و همچنین حذف اثرات نور محیط در طیف گیری انجام شد. برای این کار از یک شاهد سفید با بازتاب بیشینه برای گرفتن طیف شاهد استفاده شد و برای طیف تاریک نیز جلوی درپوش فیبر نوری دریافت کننده ی نور بسته شد. نمونه ها در زیر آشکارساز و

1. Linear Discriminant Analysis
2. Principal Component Analysis
3. Support Vector Machine
4. Full Cross-Validation

۳- نتایج و بحث

۳-۱- طبقه بندی نمونه ها به روش PCA

شکل ۱، طیف‌های جذبی NIR نمونه‌ها را در محدوده طول موج ۳۵۰ تا ۲۵۰۰ نانومتر نشان می‌دهد. مطابق این شکل، طیف‌ها یک روند افزایشی و سه پیک جذبی مشخص در اطراف طول موج های ۱۲۱۴، ۱۴۶۲ و ۱۹۶۶ نانومتر داشتند که با توجه به چگونگی توزیع اورتون‌های پیوندهای اصلی، می‌توان آنها را به ترتیب ناشی از اورتون‌های دوم O-H و یا NH₂، اورتون های دوم C-H، CH₂ و یا CH₃ و اورتون های اول O-H و یا NH₂ دانست [۴۶، ۴۵]. همان‌گونه که در شکل ۱ مشاهده می‌شود، ۶ پیک مهم در طیف جذبی لیموترش وجود دارد. نخستین پیک در محدوده طول موج ۴۹۰-۵۱۰ نانومتر ظاهر شده است. کلروفیل‌ها یا رنگدانه‌های سبز (کلروفیل a و b) در میوه‌های نارس بیشتر هستند و به تدریج با رسیدن میوه و کاهش کلروفیل‌ها، کاروتنوئیدهای زرد و سرخ که معیاری از رنگ ویژه میوه هستند، افزایش می‌یابند. از این رو و با توجه به نواحی جذب این رنگدانه‌ها دو پیک جذبی ایجاد شده در اطراف طول موج‌های ۴۹۰-۵۱۰ نانومتر و ۶۳۰-۶۸۰ نانومتر به ترتیب به جذب کاروتنوئیدها و کلروفیل a نسبت داده شده است. محدوده ۶۳۰-۶۸۰ نانومتر در ناحیه قرمز طیف مرئی است و پیک موجود در این محدوده به دلیل جذب بالای رنگدانه‌های قرمز توسط کلروفیل بوده و نشانی از رنگ میوه است. بعد از پیک در محدوده ۶۳۰-۶۸۰ نانومتر، کاهش سریعی در شدت جذب رخ می‌دهد که در این ناحیه طیف وارد محدوده NIR می‌شود. پیک بعدی در طول موج حدود ۹۷۰ نانومتر رخ می‌دهد که به جذب آب و کربوهیدرات‌ها مربوط است. در واقع وقوع پیک در این محدوده را می‌توان به اورتون سوم پیوندهای O-H، و یا اورتون دوم N-H نسبت داد. به دلیل اینکه میوه‌ها بیش از ۸۰٪ رطوبت دارند، وجود چنین پیکی در این محدوده قابل انتظار است. پیک چهارم در محدوده طول موج ۱۱۹۰-۱۱۵۰ نانومتر رخ داده است که به اورتون‌های دوم پیوندهای C-H مربوط می‌شود. پیک پنجم در طول موج حدودی ۱۴۵۰ نانومتر مشاهده می‌شود که این پیک نیز به اورتون‌های دوم C-H، H₂O، ترکیبات آروماتیک و اسیدهای ارگانیک نسبت داده می‌شود. در نهایت وقوع پیک ششم در اطراف ۱۹۵۰ نانومتر را

مدل‌های آماری مورد استفاده قرار می‌گیرد. به عبارت دیگر پس از ساخت مدل در مرحله واسنجی^۵، با ارزیابی مجدد مدل، به بررسی اعتبار و پایداری مدل نیز پرداخته شد.

۳-۲-۱- تحلیل تفکیک خطی (LDA)

این روش ترکیبی خطی از تمام ویژگی‌هایی که موجب ایجاد کلاسه بندی در یک سری نمونه می‌گردد، ایجاد می‌کند. این تابع نسبت واریانس بین گروهی را به واریانس درون گروهی افزایش می‌دهد. تبدیلات و انتقالات در این تابع به صورتی انجام می‌شود که وقتی مشاهدات جدیدی وارد می‌شوند، جهت پیش بینی اختلافات بین گروه‌ها به حداکثر مقدار خود می‌رسد [۴۴].

۳-۲-۲- تحلیل مؤلفه های اصلی (PCA)

روشی آماری که از انتقال متعامد، برای تبدیل مجموعه ای از متغیرهای مشاهده شده دارای همبستگی به مجموعه ای از متغیرهای غیر همبسته خطی که اجزای اصلی هستند، استفاده می‌کند. این تبدیل بدین صورت است که اولین جزء اصلی دارای واریانس بالایی است و سپس اجزای دیگر نیز دارای واریانس بالا و البته با محدودیت می‌باشند که در تعامد با اجزای پیشین می‌باشند. روش PCA حساس و با دقت بالا در یافتن متغیرهای اصلی می‌باشد. PCA از جمله روش‌های متداول در تجزیه و تحلیل داده‌ها و کاهش ابعاد در سیستم‌های چند متغیره است [۴۴]. از این روش در تشخیص الگو جهت تمایز بین نمونه‌های لیموترش طی روزهای مختلف استفاده شد.

۳-۳-۲- ماشین بردار پشتیبان (SVM)

هدف از این روش یافتن صفحه‌ای بهینه با کمترین فاصله بین تمام داده‌ها و نقاط می‌باشد. الگوریتم آموزشی SVM، مدلی جدید به دسته داده‌ها نسبت می‌دهد یا داده‌ها را به یک طبقه بندی خطی مرزی غیر احتمالی تبدیل می‌کند. نتیجه این مدل، نمایشی از داده‌ها در فضای چند بعدی است که داده‌ها در طبقه‌هایی تقسیم بندی شده‌اند و توسط یک صفحه محدودده‌های بین داده‌ها مشخص می‌گردد. داده‌های جدید نیز در همان فضا قرار می‌گیرند و بر اساس محدوده‌ای که روی صفحه قرار می‌گیرند، طبقه بندی آنها پیش بینی می‌شود [۴۴].

در این حالت میزان ضریب همبستگی و خطا به ترتیب برابر با ۰/۹۱۸ و ۰/۴۸ درجه بریکس تعیین گردید. گومز و همکاران (۲۰۰۵) از طیف سنجی مرئی و مادون قرمز نزدیک برای پیشگویی خصوصیات کیفی میوه نارنگی استفاده نمودند و توانستند خصوصیات فیزیولوژیکی مهم میوه مانند مواد جامد محلول، pH و سفتی بافت را با داده های بدست آمده از طیف سنجی پیشگویی کنند. در همین راستا و با استفاده از روش های PLS و PCR میزان مواد جامد محلول را با ضریب همبستگی ۰/۹۴ و میزان خطای ۰/۳۲۵ درجه بریکس پیشگویی نمودند. در پژوهشی دیگر باقرپور و همکاران (۲۰۱۵) مقدار SSC و SC را در ریشه های چغندر قند اندازه گیری نمودند. نتایج حاصل از این پژوهش نشان داد که بکارگیری روش NIR در محدوده ی طول موج ۱۶۰۰-۹۰۰ نانومتر توانایی خوبی در تشخیص مقدار SSC چغندر قند دارد ولی در تشخیص مقدار ساکارز تا حدودی ضعیف تر عمل می کند که این نتیجه ممکن است به علت غالب بودن پیوندهای جذبی O-H آب در این محدوده باشد. اسچمیلویتچ و همکاران (۱۹۹۹) از یک سیستم نیمه خودکار که بر مبنای طیف سنجی NIR در دامنه ۱۲۰۰-۲۴۰۰ نانومتر کار می کرد، برای تخمین SSC و رطوبت خرمای حیانی استفاده کردند، نتایج این تحقیق یک خطای استاندارد پیشگویی (SEP) ۱٪ و ضریب تبیین (R^2) ۰/۸۱ در هر دو مدل تخمین SSC و رطوبت را نشان دادند.

می توان به اورتون اول O-H و اسیدهای ارگانیک نسبت داد. با توجه به شکل ۲، طیف های جذبی NIR نمونه های لیموترش در طی ۴ روز مختلف قابل مشاهده می باشد. در بخش اول گراف، اطلاعات خاصی استخراج نمی شود زیرا طیف ها بسیار شبیه هم بوده و بر هم منطبق شده اند. در طول موج های ۳۵۰-۶۹۰ نانومتر طیف ها از هم تفکیک شده اند و قابل تشخیص می باشند که علت آن است، طیف های جذبی یا عبوری محصولات کشاورزی معمولا در نواحی کمتر از ۷۰۰ نانومتر بیشتر متأثر از رنگ می باشند و چون این پژوهش در فواصل زمانی مختلف سنجش شده است و رنگ محصول در طی زمان تغییر کرده پس پاسخ طیفی آنها مختلف است. در طول موج های ۲۵۰۰-۱۳۷۰ نانومتر تفکیک طیف ها بصورت کاملا مشهود می باشد [۴۷].

شائو و همکاران (۲۰۰۶) با استفاده از طیف سنجی بازتابشی مادون قرمز نزدیک به بررسی خصوصیات کیفی گوجه فرنگی مانند سفتی (شامل نیروی فشاری و نیروی لازم برای سوراخ کردن) و همچنین مواد جامد محلول و pH پرداختند و توانستند این خصوصیات را با ضریب همبستگی بالا (۰/۸۹) به صورت غیر مخرب پیشگویی کنند. در پژوهشی دیگر های کینگ و همکاران (۲۰۰۷) به اندازه گیری مواد جامد محلول در میوه هندوانه به کمک روش طیف سنجی مادون قرمز پرداختند. مدل ها به روش های کالیبراسیون کمترین توان های دوم جزئی و رگرسیون مؤلفه های اصلی تدوین شد و بهترین ضریب همبستگی با مدل PLS و پیش پردازش مشتق اول بدست آمد

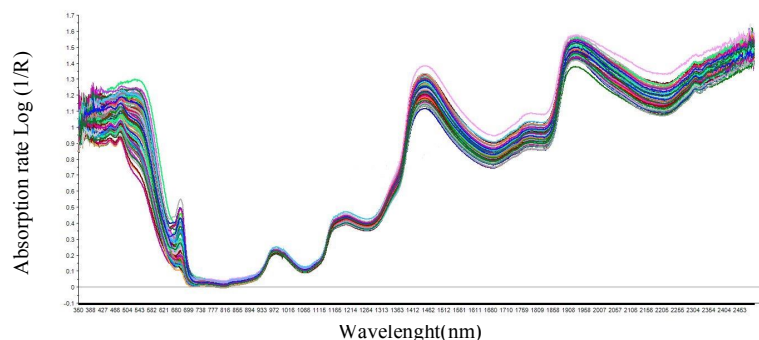


Fig 1 Absorption spectra of NIR samples of lemon

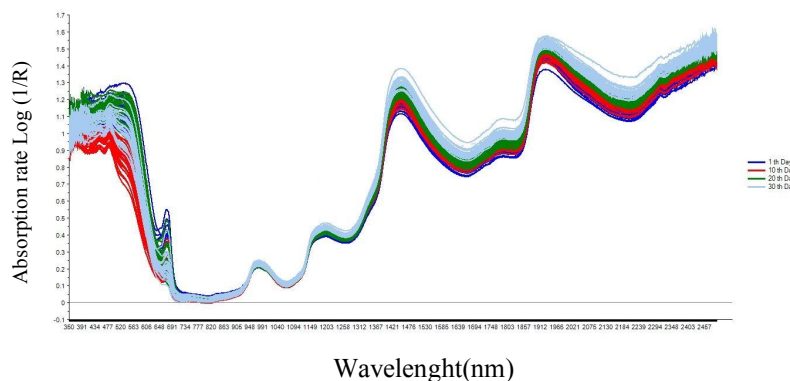


Fig 2 Absorption spectra of NIR samples of lemon species in 4 different days

متغیرهای مستقل داشته‌اند که در بین آن‌ها فاکتور اول بیشترین سهم را در توصیف داده‌ها داشته است. با افزایش تعداد فاکتور دقت مدل افزایش یافته و از فاکتور ششم به بعد، افزایش تعداد مؤلفه‌های اصلی بی‌تاثیر بوده و ممکن است باعث کاهش میزان اعتبار مدل شود. عملکرد روش PCA در دسته‌بندی نمونه‌های لیموترش در شکل ۴ نشان داده شده است. همانطور که مشاهده می‌شود، دو مؤلفه PC1 و PC2 توانایی توصیف به ترتیب ۵۰ و ۴۲ درصد واریانس بین نمونه‌ها و در مجموع ۹۲٪ واریانس کل داده‌ها را داشتند.

شکل ۳، مقادیر توصیف شده متغیرهای مستقل (مقادیر جذبی در طول موج‌های مختلف) را نشان می‌دهد. مدل ساخته شده توانسته است ۹۸٪ متغیرهای مستقل را در ۶ فاکتور برآورد کند. هرچه میزان توصیف متغیر در هر یک از فاکتورها بیشتر و تعداد فاکتورهای مورد استفاده کمتر باشد، مدل دارای پایداری بیشتری خواهد بود. افزایش شیب خطوط بین متغیرهای پنهان نشان از توانایی بیشتر آن مؤلفه در توصیف داده‌ها است. همانطور که مشخص است، مقدار بهینه فاکتور برای مدل ۶ است و مؤلفه‌های اول، دوم و سوم توانایی بیشتری در توصیف

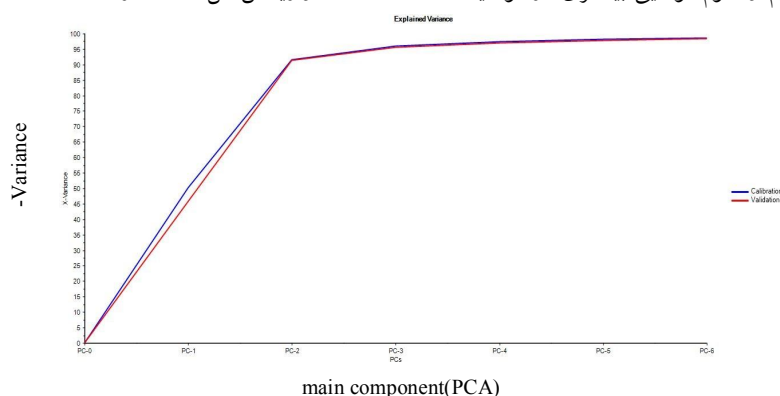


Fig 3 Described variances and factors used in PCA method in the classification of research treatments

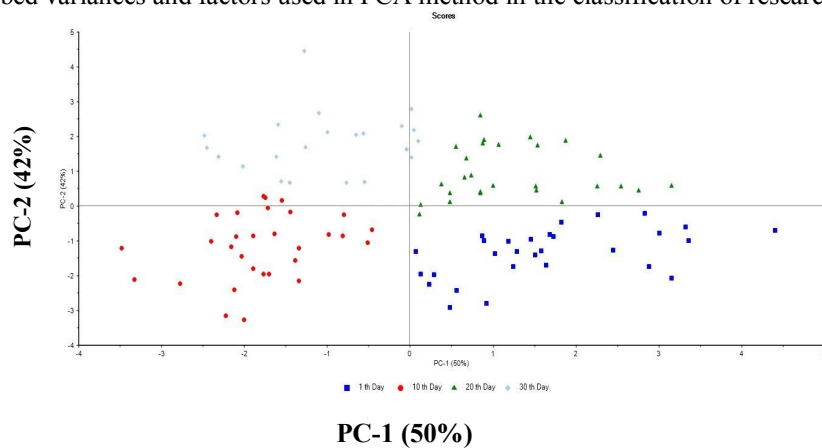


Fig 4 Performance of PCA method in classifying lemong species during 4 different days

این نوع رویکرد شامل به حداکثر رساندن نسبت واریانس بین کلاس ها به واریانس درون کلاس ها است. هدف اصلی این است که این نسبت به حداکثر رسانده شده بطوری که جدایی کافی کلاس به دست آید. رویکرد نوع خاص کلاس شامل استفاده از دو معیار بهینه برای تبدیل مجموعه داده ها بطور مستقل است.

(ب) انتقال کلاس - مستقل^۷

این رویکرد شامل به حداکثر رساندن نسبت واریانس کلی به واریانس درون کلاس است. این روش تنها از یک معیار بهینه برای تبدیل مجموعه داده ها استفاده می کند و از این رو تمام داده ها صرف نظر از هویت کلاس آنها با استفاده از این تبدیل منتقل می شوند. در این نوع LDA، هر کلاس به عنوان کلاس جداگانه در مقابل تمام کلاس های دیگر در نظر گرفته می شود [۵۴].

نتایج حاصل از روش LDA در تفکیک تیمارهای مختلف تحقیق به همراه دقت هر کدام در جدول ۱ نشان داده شده است. برای ایجاد مدل دسته بندی به روش LDA از سه نوع تابع خطی، درجه دوم و Mahalanobis همراه با ۴ نوع ضرایب وزنی (Block weighting, Constant, 1/SDev) و (Block weighting + SDev) مختلف استفاده گردید.

با توجه به جدول ۱ مدلسازی با استفاده از تابع درجه دوم در اکثر موارد منجر به ایجاد مدلی با دقت بهتری نسبت به زمانی بود که از توابع خطی و Mahalanobis برای ساخت مدل به کار گرفته شد. همچنین نتایج بدست آمده نشان داد که تابع درجه دوم، در شرایطی که ضرایب وزنی در هر سه روش یکسان باشد، مدلی با دقت بهتر نسبت به سایر توابع ایجاد میکند. در کل این روش قادر بود که گروه های مختلف تحقیق را بدون خطا و با دقت ۱۰۰٪ دسته بندی کند. از دیگر نتایجی که در جدول ۱ می توان به آن اشاره کرد، تاثیر ضرایب وزنی بر دقت روش LDA است.

همچنین نتایج نشان می دهد که این روش از قابلیت بسیار خوبی در دسته بندی نمونه های لیموترش برخوردار است و توانسته بین گروه های مختلف تمایز واضحی را ایجاد کند. در تحقیقی درصدهای مختلف فرمالین، پراکسید هیدروژن و هیپوکلریت سدیم در شیرخام با استفاده از PCA شناسایی و مورد طبقه بندی قرار گرفتند. در همین راستا و با استفاده از همین روش ۸۳٪ واریانس کل داده ها پوشش داده شدند [۵۲].

۲-۳- طبقه بندی نمونه ها به روش LDA

تحلیل تفکیک خطی یکی از روش های رایج و موفق در کاهش بعد و طبقه بندی داده از فضایی با بُعد بالا به یک فضایی با بُعد کم است. در LDA کلاسیک نتیجه مطلوب در دستیابی به حداکثر تفکیک بین گروه های مختلف با به حداقل رساندن واریانس درون کلاس و به حداکثر رساندن واریانس بین طبقه های مختلف بطور همزمان، بدست می آید. LDA اطلاعات مربوط به هر دو توزیع در کلاس و همچنین بین کلاس ها را در نظر می گیرد. در روش LDA مطلوب آنست که مدلی ساخته شود تا واریانس داده ها در داخل هر گروه به حداقل برسد و از سویی دیگر واریانس بین گروه های مختلف افزایش یابد. به عبارت دیگر، در روش LDA، سعی بر آنست که فاصله داده های یک گروه را نسبت به هم به حداقل برساند و آنها را به یکدیگر نزدیک تر کند. از سوی دیگر، داده های گروه های مختلف از یکدیگر، در فاصله حداکثری قرار داشته باشند. [۵۳]. در این روش هدف ساخت صفحه ای چند وجهی در فضای داده ها است به گونه ای که این صفحه قادر به تفکیک گروه ها باشد. بهترین صفحه چند وجهی زمانی ساخته می شود که کوواریانس بین گروه ها یکسان باشد. از طرفی دیگر، روش LDA زمانیکه تعداد داده ها زیاد باشد، عملکرد ضعیفتری از خود نشان می دهد. در روش LDA مجموعه داده ها می توانند منتقل شده و بردارهای آزمون می توانند در فضای منتقل شده با دو روش مختلف طبقه بندی شوند.

(الف) انتقال کلاس - وابسته^۶

6. Class-dependent transformation

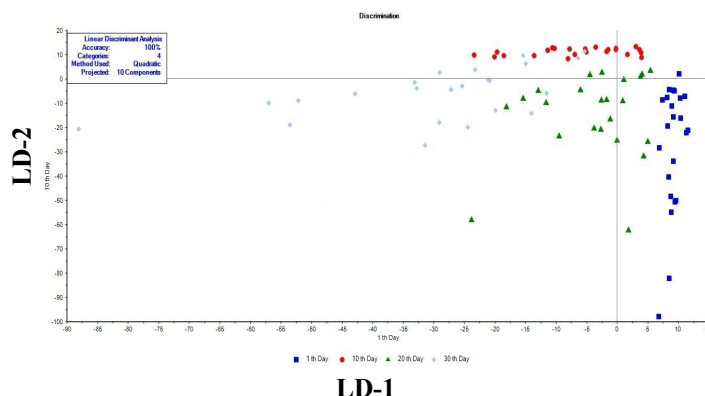
7. Class-independent transformation

Table 1 Results of LDA model of lemon samples

Accuracy (%)	Weighting coefficients	Function type	Row
97.83	1/SD _{ev}	Linear	1
98.91	1/SD _{ev}	Quadratic degree	2
98.91	1/SD _{ev}	Mahalanobis	3
97.83	Constant	Linear	4
100	Constant	Quadratic degree	5
97.83	Constant	Mahalanobis	6
97.83	Block Weighting	Linear	7
100	Block Weighting	Quadratic degree	8
98.91	Block Weighting	Mahalanobis	9
97.83	Block Weighting + SD _{ev}	Linear	10
97.91	Block Weighting + SD _{ev}	Quadratic degree	11
97.83	Block Weighting + SD _{ev}	Mahalanobis	12

به عبارت دیگر، با انتخاب ضرایب وزنی مناسب، می‌توان دقت مدل را بهبود بخشید. انتخاب ضریب وزنی مناسب به نوع داده‌ها بستگی دارد و معمولاً با آزمون و خطا می‌توان به روش مناسب اعمال ضریب وزنی دست یافت. شکل ۵ عملکرد روش LDA با استفاده از تابع درجه دوم و با ضریب وزنی ثابت (عدد ۱) در تفکیک تیمارهای تحقیق را نشان می‌دهد.

به عبارت دیگر، با انتخاب ضرایب وزنی مناسب، می‌توان دقت مدل را بهبود بخشید. انتخاب ضریب وزنی مناسب به نوع داده‌ها بستگی دارد و معمولاً با آزمون و خطا می‌توان به روش مناسب اعمال ضریب وزنی دست یافت. شکل ۵ عملکرد روش LDA با استفاده از تابع درجه دوم و با ضریب وزنی ثابت (عدد ۱) در تفکیک تیمارهای تحقیق را نشان می‌دهد.

**Fig 5** Performance of LDA method using quadratic function in separating research treatments

روش LDA قابلیت بالایی (۷۹/۰۷ درصد) در طبقه بندی الگوهای بوی تقلاب داشت. نمونه های رب خالص و تقلاب ۲۰ درصد به خوبی تفکیک شدند. نمونه ها با تقلاب ۱۰ و ۱۵ درصد با هم، همپوشانی نزدیکی داشتند، همچنین تقلاب ۵ درصد در چند مورد با نمونه های ۱۰ و ۱۵ درصد همپوشانی کمی داشتند. با افزودن حتی مقدار اندکی از یک ماده غذایی دیگر به یک ماده، الگوی بو نسبت به حالت طبیعی آن تغییر کرده و با توجه به حساسیت بالای حسگر و توان تفکیک پذیری بالای آنها، حسگر توانایی ایجاد تفاوت بین مواد فرار روی سطح حسگر را نشان می دهد [۵۶].

۳-۳- طبقه بندی نمونه ها به روش SVM

در روش SVM نیز همانند روش LDA، از صفحات چندوجهی برای تفکیک و تمایز بین گروه‌ها استفاده می‌شود با

نقاط نشان داده شده در شکل‌ها میزان پراکندگی‌ها طی ۴ روز مختلف می‌باشند و نکته حائز اهمیت آن است که داده‌ها در بخش واسنجی و اعتبارسنجی دقیقاً منطبق بر یکدیگر می‌باشند. یعنی زیر هر نقطه (داده واسنجی) داخل شکل، داده اعتبارسنجی قرار گرفته که نشان از اعتبار و پایداری بسیار خوب مدل است. در تحقیقی درصدهای مختلف وجود فرمالین، پراکسید هیدروژن و هیپوکلریت سدیم در شیرخام با استفاده از روش LDA شناسایی و طبقه بندی گردیدند و این روش به ترتیب ۷۹/۱۶، ۷۰/۸۳ و ۶۶/۶۶ درصد واریانس کل داده‌ها را پوشش داد [۵۲]. قاسمی ورنامخواستی و همکاران نیز نتایج مشابهی برای طبقه بندی زیره گزارش کردند [۵۵]. در پژوهشی که توسط محمدرزاداری و همکاران جهت بررسی تقلاب در رب گوجه فرنگی صورت پذیرفت، نشان داد که

شعاعی و سیگموئید استفاده شد. در هر دو روش C-SVM و Nu-SVM تابع خطی و درجه دوم بالاترین دقت را در طبقه بندی لیموترش ها به خود اختصاص دادند (شکل ۶، ۷، ۸ و ۹). نتایج حاصل از روش SVM در جدول ۲ نشان داده شده است. توحیدی و همکاران درصدهای مختلف وجود فرمالین، پراکسید هیدروژن و هیپوکلریت سدیم در شیر با دقت ۹۴/۶۴، ۹۲/۸۵ و ۸۷/۷۵ درصد با کمک SVM طبقه بندی کردند [۵۲]. در پژوهشی دیگر با کمک روش Nu-SVM، تابع پایه شعاعی با دقت ۹۸/۸۴ و ۸۸/۱۴ درصد به ترتیب برای آموزش و اعتبارسنجی بیشترین دقت طبقه بندی برای نمونه های رب گوجه فرنگی را به خود اختصاص دادند [۵۶].

این تفاوت که شرط یکسان بودن کوواریانس بین گروه‌ها مورد استفاده قرار نمی‌گیرد. در این روش، صفحه چندوجهی با در نظر گرفتن یک حاشیه اطمینان محاسبه شده و سعی می‌شود که جهت و محل قرارگیری آن به گونه‌ای باشد که حاشیه در نظر گرفته شده، حداکثر باشد. روش SVM در ایجاد مدل دسته‌بندی برای مواقعی که ابعاد ماتریس داده‌ها بزرگ باشد، عملکرد بسیار بهتری از خود نشان می‌دهد [۵۴].

برای طبقه بندی نمونه‌ها در روش SVM از دو نوع C-SVM و Nu-SVM استفاده گردید. پارامترهای C، Nu و γ با سعی و خطا و از طریق به حداقل رساندن خطای مدل در بخش واسنجی با بخش اعتبارسنجی انتخاب شدند. در ادامه از چهار نوع تابع کرنل شامل خطی، درجه دوم، تابع پایه

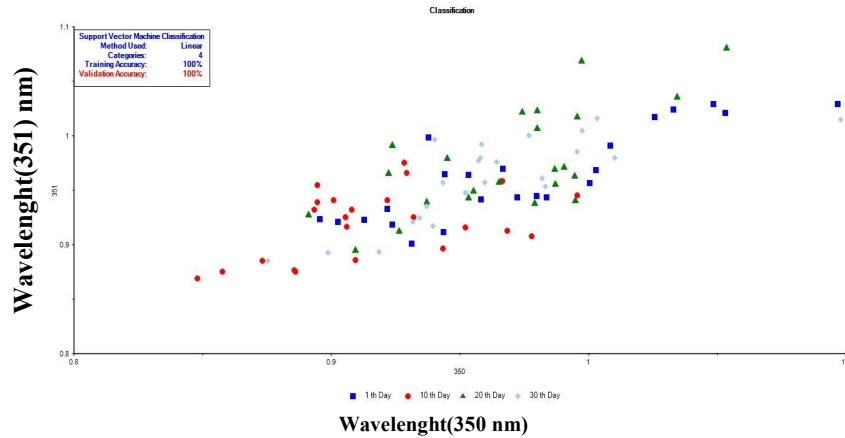


Fig 6 C-SVM linear function with 100% accuracy

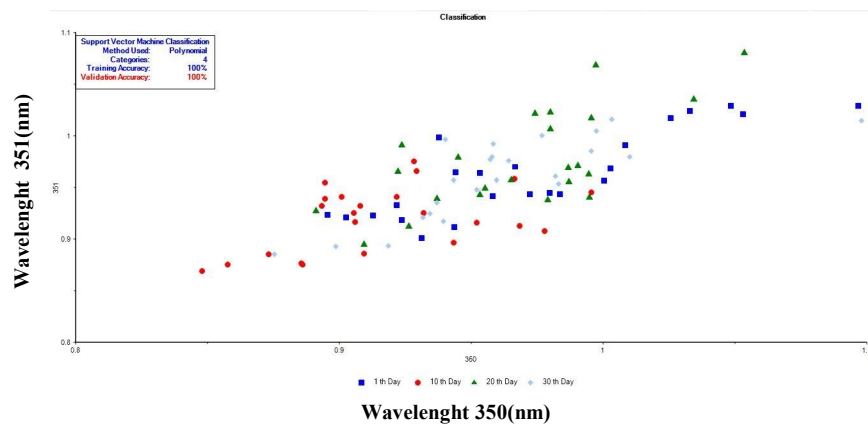


Fig 7 Quadratic C-SVM function with 100% accuracy

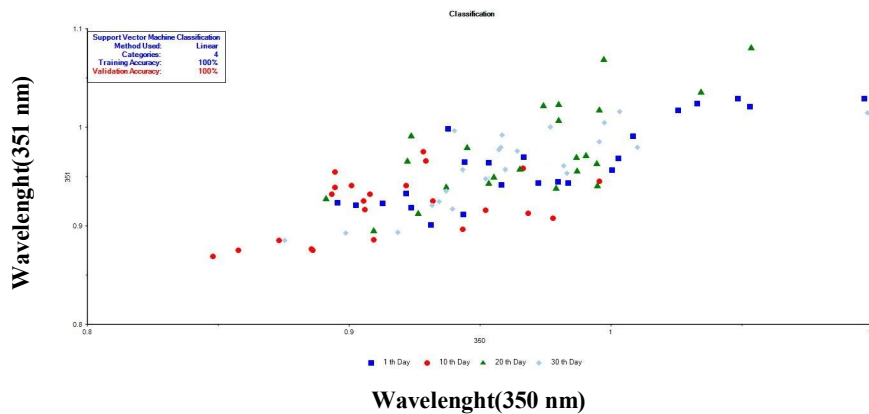


Fig 8 Nu-SVM linear function with 100% accuracy

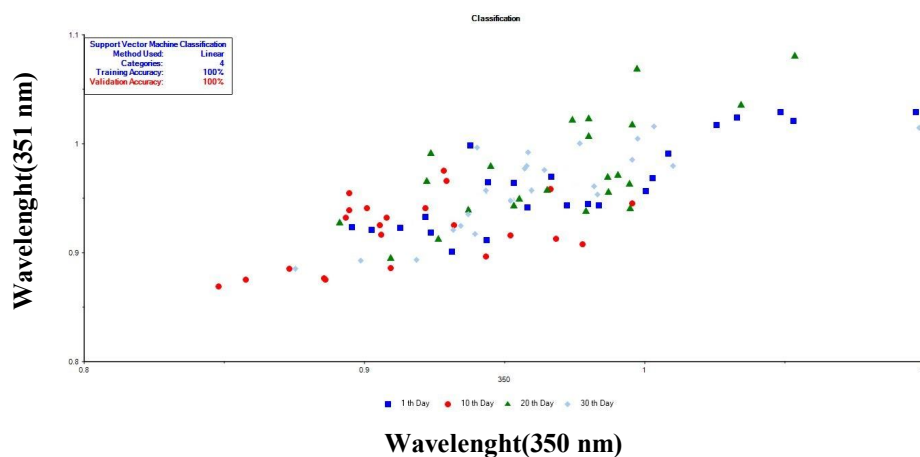


Fig 9 Second-degree Nu-SVM function with 100% accuracy

ناشناخته x ، اگر در معادله $w^T x + b \geq 1$ برقرار باشد نمونه متعلق به اولین کلاس و در صورتی که $w^T x + b \leq -1$ باشد نمونه متعلق به کلاس دوم می باشد.

در صورتیکه داده ها به صورت خطی تفکیک پذیر نباشند از یک تابع هسته استفاده می شود که داده ها را به یک فضای ویژگی دیگر با قابلیت تفکیک خطی می برد (شکل ۱۰).

روش بحث شده برای ماشین بردار پشتیبان برای مسأله طبقه بندی دو کلاسه بود. ماشین بردار پشتیبان برای مسائل چند کلاسی نیز قابل تعمیم است. معمولاً به این منظور از دو روش "یکی در برابر بقیه" و "یک در برابر یک" استفاده می شود. در روش "یکی در برابر بقیه" برای هر کلاس یک ماشین بردار پشتیبان ساخته می شود که آن کلاس را از بقیه (M-1) کلاس جدا می کند.

رویکرد SVM به این صورت است که در فاز آموزش، سعی می شود که مرز تصمیم گیری به گونه ای انتخاب گردد که حداقل فاصله آن با هر یک از دسته های مورد نظر ماکزیمم گردد. این نوع انتخاب باعث می شود که تصمیم گیری ما در عمل، شرایط نویزی را به خوبی تحمل کند و پاسخ دهی خوبی داشته باشد. این نحوه انتخاب مرز بر اساس نقاطی به نام بردارهای پشتیبان انجام می شود. فرض کنید که کلاس ها به صورت خطی قابل تفکیک شدن هستند از این رو یک صفحه (یا یک خط در حالت دو بعدی) می تواند این دو را از هم جدا کند. معادله کلی این صفحه به صورت زیر است:

$$w^T x + b = 0 \quad (3)$$

بنابراین $w^T x + b = +1$ صفحه ای (صفحه مثبت) است که بردارهای پشتیبان در کلاس اول هستند و $w^T x + b = -1$ (صفحه منفی) صفحه ی مربوط به بردارهای پشتیبان در کلاس دیگر می باشد. برای نمونه

- 8. Plus Class
- 9. Minus Class

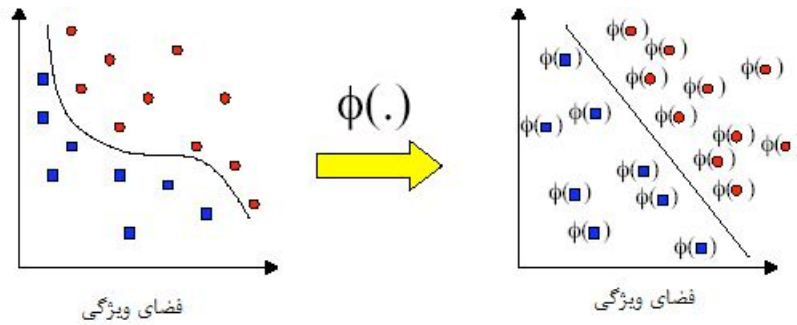


Fig 10 Use a kernel function to move data to a linear attribute space

استفاده در این روش برابر $M(M-1)/2$ است. برای تشخیص تعلق یک نمونه به یک کلاس، از رأی گیری استفاده می شود [۵۴].

تعداد SVM ها در این روش برابر M است. یک نمونه، به کلاسی که دارای بیشترین مقدار دقت برای جداکنندگی باشد تعلق می گیرد. در روش "یک در برابر یک" SVM، برای هر جفتی از کلاس ها اجراء می شود. تعداد SVM های مورد

Table 2 Results of Nu-SVM and C-SVM models under kernel function

Nu-SVM				C-SVM				Kernel Function
Validation	Calibration	γ	Nu	Validation	Calibration	γ	C	
100	100	1	0.01	100	100	1	100	Linear
100	100	0.01	0.01	100	100	100	0.01	
97.82	100	0.01	0.01	100	100	0.02	100	Quadratic degree
90.21	93.47	1×10^{10}	0.8	100	100	5.9×10^{10}	2.78×10^{10}	

نتیجه را حاصل کرده و با مدل LDA مطابقت دارد. این تطابق نتایج نشان می دهد که داده ها بطوری توزیع شده اند که منحنی های درجه دوم برای جداسازی گروه ها از هم دقت، توانایی و کارایی بهتری از خود نشان می دهند.

۴- نتیجه گیری

در این پژوهش از یک سامانه طیف سنجی مرئی / فروسرخ نزدیک (Vis/NIR) به منظور طبقه بندی نمونه های لیموترش گونه کی لایم در خلال دوره انبارمانی با روش های شناسایی الگو شامل PCA، LDA و SVM مورد بررسی قرار گرفتند. نتایج بدست آمده نشان داد که طیف سنجی (Vis/NIR) قادر به تفکیک نمونه های لیموترش بر اساس مدت زمان نگهداری در انبار است. هرچند روش های PCA، LDA و SVM توانستند با دقت خوبی نمونه های لیموترش را دسته بندی کنند، اما دو روش LDA و SVM با دقت ۱۰۰٪ بهترین روش برای دسته بندی نمونه ها بودند. همچنین، بر اساس نتایج حاصله، تابع درجه ۲، بهترین تابع برای ساخت مدل های دسته بندی به روش LDA و SVM تعیین گردید.

مطابق جدول ۲ و نتایج حاصل از مدل سازی به روش SVM، تابع درجه دوم در هر دو مدل C-SVM و Nu-SVM دارای دقت ۱۰۰ درصدی بوده و برای جداسازی و طبقه بندی داده های آزمون بسیار کاربردی می باشد. بر اساس داده های جدول ۲، می توان دریافت که روش C-SVM نسبت به Nu-SVM از توانایی بیشتری در مدل سازی دسته بندی داده های تحقیق دارد. زیرا در روش C-SVM در تمامی توابع مورد استفاده، دقت مدل در هر دو بخش واسنجی و اعتبارسنجی ۱۰۰٪ بوده است. این در حالی است که برای Nu-SVM، فقط در توابع خطی و درجه دوم دقت مدل در هر دو بخش واسنجی و اعتبارسنجی ۱۰۰٪ بوده و زمانی که توابع شعاعی و سیگموئید برای ساخت مدل مورد استفاده قرار گرفت، دقت مدل به ویژه در بخش اعتبارسنجی کاهش یافته است. لازم به ذکر است که هر دو روش مدل سازی SVM و LDA در تفکیک تشخیص و دسته بندی تیمارهای تحقیق عملکرد بسیار قابل قبولی داشتند. بطوریکه هر دو روش با استفاده از تابع درجه دوم توانستند با حداکثر دقت (۱۰۰٪) موفق به طبقه بندی داده ها شوند. بطور کلی با بررسی مدل SVM به این نتیجه رسیدیم که تابع منحنی درجه دوم برای داده های حاصل از آزمون، بهترین

۵- منابع

- [11] Steensland, A., & Zeigler, M. (2021). Productivity in Agriculture for a Sustainable Future. In: Campos H, Editor. The Innovation Revolution in Agriculture. Springer, 33–69. Doi.org/443.webvpn.fjmu.edu.cn/10.1007/978-3-030-50991-0_2
- [12] Moghimi, A., Aghkhani, M. H., Sazgarnia, A., & Sarmad, M. (2010). Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. *Journal of Biosystems Engineering*, 106, 295-302. Doi.org/10.1016/j.biosystemseng.2010.04.002
- [13] Liu, Y.d., Ying, Y.b., Fu, X., & Lu, H. (2006). Experiments on predicting sugar content in apples by FT-NIR technique. *Journal of Food Engineering*, 80: 986-989. Doi.org/10.1016/j.jfoodeng.2006.06.035
- [14] Flores, K., Sánchez, M.T., Pérez-Marín, D., Guerrero, J.E., & Garrido-Varo, A. (2008). Feasibility in NIRS instruments for predicting internal quality in intact tomato. *Journal of Food Engineering*, 91: 311-318. Doi.org/10.1016/j.jfoodeng.2008.09.013
- [15] Liu, Y., Sun, X., & Ouyang, A. (2009). Nondestructive measurements of soluble solid content of navel orange fruit by visible-NIR spectrometric technique with PLS and PCABPNN. *Journal of LWT Food Science and Technology*, 43: 602–607. Doi.org/10.1016/j.lwt.2009.10.008
- [16] Paz, P., Sánchez, M.T., Pérez-Marín, D., Guerrero, J.E., & Garrido-Varo, A. (2009). Instantaneous quantitative and qualitative assessment of pear quality using near infrared spectroscopy. *Journal of Computers and electronics in agriculture*, 69: 24-32. Doi.org/10.1016/j.compag.2009.06.008
- [17] Jamshidi, B., Minaei, S., Mohajerani, E., & Ghassemian, H. (2012). Reflectance Vis/NIR spectroscopy for nondestructive taste characterization of Valencia oranges. *Journal of Computers and Electronics in Agriculture*, 85: 64-69. Doi.org/10.1016/j.compag.2012.03.008
- [18] Walsh, K.B., Blasco, J., Zude-Sasse, M., & Sun, X. (2020). Visible-NIR 'point' spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. *Journal of Postharvest Biology and Technology*, 168, 111246. Doi.org/10.1016/j.postharvbio.2020.111246
- [1] Food and Agriculture Organization (FAO), 2017., Available at: www.fao.org/publications [Accessed: 2020-10-20]
- [2] Liu, Y., Heying, E., & Tanumihardjo, S.A. (2012). History, Global Distribution, and Nutritional Importance of Citrus Fruits. *Journal of Food Science and Food Safety*, 11(6), 530–545. Doi.org/10.1111/j.1541-4337.2012.00201.x
- [3] Codex Alimentarius Commission (CAC), (2020)., Available at: http://www.fao.org/faowho-codexalimentarius/codex-texts/list-standards/En/[Accessed: 2020-12-07]
- [4] Ladaniya, M.S. (2008). Citrus fruit: Biology, Technology and Evaluation. 1st Edition. Academic Press (Elsevier); San Diego. 576 p
- [5] Food and Agriculture Organization (FAO), 2010., Available at: www.fao.org/statistics.htm
- [6] Rozbahani, A., Movahhed, S., & Ahmadi Chenarbon, H. (2019). Mathematical modeling of hydrodynamic properties of lime (Mexican lime). *Journal of Food Process Engineering*. 42(4), e13054. doi.org/10.1111/jfpe.13054
- [7] Rivera Cabrera, F., Artes Hernandez, F., & Kader, A.A. (2006). Quality retention and potential shelf-life of fresh-cut lemons as affected by cut type and temperature. *Journal of Postharvest Biology and Technology*, 43 (2): 245-254. Doi.org/10.1016/j.postharvbio.2006.09.009
- [8] Conrad, Z., Niles, M.T., Neher, D.A, Roy, E.D, Tichenor, N.E, & Jahns, L. (2018). Relationship between food waste, diet quality, and environmental sustainability. *Journal of PLoS ONE*, 13(4), e0195405. Doi.org/10.1371/journal.pone.0195405
- [9] Yahaya, S.M., & Mardiyya, A.Y. (2019). Review of Post-Harvest Losses of Fruits and Vegetables. *Biomedical Journal of Scientific and Technical Research*, 13(4), 10192–10200. Doi.org/10.26717/BJSTR.2019.13.002448
- [10] Food and Agriculture Organization (FAO), 2015., Available at: www.fao.org/3/a-i4951e.pdf [Accessed: 2020-12-10]

- and Technology, 50(1), 31–36. Doi.org/10.1016/j.postharvbio.2008.03.013
- [27] Tian, X., Wang, Q., Li, J., Peng, F., & Huang, W. (2018). Non-destructive prediction of soluble solids content of pear based on fruit surface feature classification and multivariate regression analysis. *Journal of Infrared Physics and Technology*, 92, 336–344. Doi.org/10.1016/j.infrared.2018.06.019
- [28] Li, Y., Jin, G., Jiang, X., Yi, S., & Tian, X. (2020). Non-destructive determination of soluble solids content using a multi-region combination model in hybrid citrus. *Journal of Infrared Physics and Technology*, 104, 103138. Doi.org/10.1016/j.aiaa.2020.05.001
- [29] Kim, K. B., Lee, S., Kim, M. S., & Cho, B.K., (2008). Determination of apple firmness by nondestructive ultrasonic measurement. *Journal of Postharvest Biology and Technology*, 52(1), 44–48. Doi.org/10.1016/j.postharvbio.2008.04.006
- [30] Aboudaoud, I., Faiz, B., Aassif, E., Izbaim, D., Abassi, D.E., & Malainine, M. (2012). Thematurity characterization of orange fruit by using high frequency ultrasonic echo pulse method. Iop Conference, 42, 012038. Doi.org/10.1088/1757-899X/42/1/012038
- [31] Blasco, J., Aleixos, N., Gomez, J., & Molto, E. (2007). Citrus sorting by identification of the most common defects using multispectral computer vision. *Journal of Food Engineering*, 83(3), 384–393. Doi.org/10.1016/j.jfoodeng.2007.03.027
- [32] Elmasry, G., Wang, N., & Vigneault, C. (2008). Detecting chilling injury in red delicious apple using hyperspectral imaging and neural networks. *Journal of Postharvest Biology and Technology*, 52 (1), 1–8. Doi.org/10.1016/j.postharvbio.2008.11.008
- [33] Lopezgarcia, F., Andreugarcia, G., Blasco, J., Aleixos, N., & Valiente, J. (2010). Automatic detection of skin defects in citrus fruits using a multivariate image analysis approach. *Journal of Computers and Electronics in Agriculture*, 71(2), 189–197. Doi.org/10.1016/j.compag.2010.02.001
- [34] Hong, X., & Wang, J. (2013). Detection of adulteration in cherry tomato juices based on electronic nose and tongue: comparison of different data fusion approaches. *Journal of Food Engineering*, 126, 89–97. Doi.org/10.1016/j.jfoodeng.2013.11.008
- [19] Walsh, K.B., McGlone, V.A., & Han, D.H. (2020). The uses of near infra-red spectroscopy in postharvest decision support: A review. *Journal of Postharvest Biology and Technology*, 163, 111139. Doi.org/10.1016/j.postharvbio.2020.111139
- [20] Alhamdan, A. M., Fickak, A., & Atia, A. R. (2019). Evaluation of sensory and texture profile analysis properties of stored Khalal Barhi dates nondestructively using Vis/NIR spectroscopy. *Journal of Food Process Engineering*, 42(6). Doi.org/10.1111/jfpe.13215
- [21] Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Journal of Postharvest Biology and Technology*, 46, 99–118. Doi.org/10.1146/annurev-food-030713-092410
- [22] Jha, S.N. (2010). *Nondestructive Evaluation of Food Quality*. Springer Heidelberg Dordrecht. London. New York
- [23] Brezmes, J., Llobet, E., Vilanova, X., Orts, J., Saiz, G., & Correig, X., (2001). Correlation between electronic nose signals and fruit quality indicators on shelf-life measurements with pink lady apples. *Journal of Sensors and Actuators B: Chemical*, 80(1), 41–50. Doi.org/10.1016/S0925-4005(01)00867-X
- [24] Saevels, S., Lammertyn, J., Berna, A.Z., Veraverbeke, E., Natale, C. D., & Nicolai, B. (2003). Electronic nose as a non-destructive tool to evaluate the optimal harvest date of apples. *Journal of Postharvest Biology and Technology*, 30(1), 3–14.
- [25] Sanaeifar, A., Mohtasebi, S.S., Ghasemi-Varnamkhasti, M., Ahmadi, H., & Lozano, J. (2014). Development and application of a new low cost electronic nose for the ripeness monitoring of banana using computational techniques (PCA, LDA, SIMCA and SVM). *Czech Journal of Food Sciences*, 32 (6), 538–548. Doi.org/doi.org/10.17221/113/2014-CJFS
- [26] Sinelli, N., Spinardi, A., Egidio, V.D., Mignani, I., & Casiraghi, E. (2008). Evaluation of quality and nutraceutical content of blueberries (*Vaccinium corymbosum* L.) by near and mid-infrared spectroscopy. *Journal of Postharvest Biology*

- Multivariate Data Analysis for Detection and Quantification of Apricot Kernel in Almond Powder. *Journal of Food Analytical Methods*, 10, 3312-3320. Doi.org/10.1007/s12161-017-0903-5
- [45] Cen, H. & He, Y. (2006). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Journal of Trends in Food Science and Technology*, 18, 72-83. Doi.org/10.1016/j.tifs.2006.09.003
- [46] Jamshidi, B., Minaei, S., Mohajerani, E. & Ghassemian, H. (2013). Linear multivariate model based on NIR spectroscopy for non-destructive internal quality prediction of orange. Proceeding of the 19th Iranian Conference on Optics and Photonics, and 5th Iranian Conference on Photonics Engineering. Jan. 22-24. Zahedan. Iran. (in Farsi)
- [47] Gomez, H., He, Y., & Pereira, A.G. (2005). Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR spectroscopy techniques. *Journal of Food Engineering*, 77:313-319. Doi.org/10.1016/j.jfoodeng.2005.06.036
- [48] Shao, Y., Gomez, H., Pereir, G., Qiu, Z., & Zhag, Y. (2006). Visible/near infrared spectrometric technique for nondestructive assessment of tomato 'Heatwave' (*Lycopersicum esculentum*) quality characteristics. *Journal of Food Engineering*, 85:672-678. Doi.org/10.1016/j.jfoodeng.2006.12.026
- [49] Hai-qing, T., Yi-bin, Y., Hui-shan, L., Xia-ping, F., & Hai-yan, Y. (2007). Measurement of soluble solids content in watermelon by Vis/NIR diffuse transmittance technique. *Journal of Zhejiang University SCIENCE B*, 8: 105-110. Doi.org/10.1631/jzus.2007.B0105
- [50] Bagherpour, H., Minaei, S., Abdollahian, N. M., & Khorasani Fardvani, M. E. (2017). Non-Destructive Determination of Sugar Content in Root Beet by Near Infrared Spectroscopy (NIRS).determination of dry matter in onions. *Journal of the American Society for Horticultural Science*, 110(2),297-303. Doi.org/10.22059/ijbse.2017.208413.664785
- [51] Schmilovitch, Z., Hoffman, A., Egozi, H., Ben-Zvi, R., Bernstein, Z., & Alchanatis, V. (1999). Maturity determination of fresh dates by near infrared spectrometry. *Journal*
- [35] Mendoza, F., Lu, R., & Cen, H. (2012). Comparison and fusion of four nondestructive sensors for predicting apple fruit firmness and soluble solids content. *Journal of Postharvest Biology and Technology*, 73, 89-98. Doi.org/10.1016/j.postharvbio.2012.05.012
- [36] Alhamdan, A. M., & Atia, A. (2017). Non-destructive method to predict Barhi dates quality at different stages of maturity utilising near-infrared (NIR) spectroscopy. *International Journal of Food Properties*, 20(sup3). Doi.prg/10.1080/10942912.2017.1387794
- [37] Wigati, L. P., Sutrisno, & Darmawati, E. (2019). Losses and waste of tomato and red chilli along the supply chain. *IOP Conference Series: Earth and Environmental Science*, 230, 012001. Doi.org/10.1088/1755-1315/230/1/012001
- [38] Sánchez, J. C. (2012). Using near-infrared spectroscopy to predict postharvest quality. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 7(021). Doi.org/10.1079/pavsnr.20127021
- [39] Kader, K.K. (2000). Quality assurance of harvested horticultural perishables. *Acta Horticulturae*, 553, 8-51.
- [40] Quilitzsch, R., Baranska, M., Schulz, H., & Hoberg, E. (2005). Fast determination of carrot quality by spectroscopy methods in the UV-VIS, NIRS and IR range. *Journal of Applied Botany and Food Quality*, 79:163-7.
- [41] Dull, G.G., Birth, G.S., & Leffler, R.G. (1989). Use of near infrared analysis for the nondestructive measurement of dry matter in potatoes. *American Potato Journal*, 66:215-25.
- [42] McGoverin, C.M., Weeranantanaphan, J., Downey, G., & Manley, M. (2010). The application of near infrared spectroscopy to the measurement of bioactive compounds in food commodities. *Journal of Near Infrared Spectroscopy*, 18:87-11.
- [43] Bureau, S., Ruiz, D., Reich, M., Gouble, B., Bertrand, D., & Renard, C.M.G.C. (2009). Rapid and non-destructive analysis of apricot fruit quality using FT-near-infrared spectroscopy. *Journal of Food Chemistry*, 113: 1323-1328. Doi.org/10.1016/j.foodchem.2008.08.066
- [44] Esteki, M., Farajmand, B., Kolahderazi, Y., & Simal-Gandara, J. (2017). Chromatographic Fingerprinting with

Modern multivariate statistical techniques, Springer

- [55] Ghasemi-Varnamkhasti, M., Tohidi, M., Mishra, P., & Izadi, Z. (2018). Temperature modulation of electronic nose combined with multi-class support vector machine classification for identifying export caraway cultivars. *Journal of Postharvest Biology and Technology*, 138, 134-139. Doi.org/10.1016/j.postharvbio.2018.01.011
- [56] Mohammad-Razdari, A., Ghasemi-Varnamkhasti, M., Yoosefian, S.H., Siadat, M., Izadi, Z., & Rostami, S. (2018). Detection of pumpkin puree adulteration in tomato paste using a gas sensor array. *Journal of New Food Technologies*, 6(1), 137-148. Doi.org/10.22104/jift.2018.2982.1726
- of the Science of Food and Agriculture*. 79(1), 86-90. doi.org/10.1002/(SICI)1097-0010(199901)79
- [52] Tohidi, M., Ghasemi-Varnamkhasti, M., Ghafarinia, V., Bonyadian, M., & Mohtasebi, S. S. (2017). Development of a metal oxide semiconductor-based artificial nose as a fast, reliable and non-expensive analytical technique for aroma profiling of milk adulteration. *International Dairy Journal*, 77, 38-46. Doi.org/10.1016/j.idairyj.2017.09.003
- [53] Heidarbeigi, K., Mohtasebi, S. S., Foroughirad, A., Ghasemi-Varnamkhasti, M., Rafiee, S., & Rezaei, K. (2014). Detection of adulteration in saffron samples using electronic nose. *International Journal of Food Properties*, 18(7), 1391-1401. Doi.org/10.1080/10942912.2014.915850
- [54] Izenman, A. J. (2013). Linear discriminant analysis. PP. 237-280. *In*.



Feasibility of Visible/Near Infrared (Vis/NIR) Spectroscopy capability in classification of lemon samples during storage period by PCA, LDA and SVM identification methods

Goodarzi, N. ¹, Movahhed, S. ^{2*}, Shakouri, M. J. ³, Ahmadi Chenarbon, H. ⁴

1. Ph.D student, Department of Food Science and Technology, College of Agriculture, North Tehran Branch, Islamic Azad University, Tehran, Iran.
2. Associated Professor, Department of Food Science and Technology, College of Agriculture, Varamin - Pishva Branch, Islamic Azad University, Varamin, Iran.
3. Assistant Professor, Department of Food Science and Technology, College of Agriculture, North Tehran Branch, Islamic Azad University, Tehran, Iran.
4. Assistant Professor, Department of Agronomy, College of Agriculture, Varamin - Pishva Branch, Islamic Azad University, Varamin, Iran.

ABSTRACT

Today, the increasing process of food waste and agricultural products is one of the serious challenges in the most countries, especially in developing countries, so one of the serious policies of governments in the food security is to reduce the waste and maintain the quality of agricultural products. So far, several methods have been used to measure the quality of agricultural products, only some of which are technically and industrially justified. Vis / NIR Spectroscopy method is one of the methods that has been considered and used in evaluating the qualitative characteristics of agricultural products due to its high speed and accuracy. In this regard, in the present study, visible/near infrared Spectroscopy was used to measure the qualitative changes and classification of K-Lime samples of lemon during the storage period (10, 20 and 30 days). In order to analyze the qualitative characteristics and classify the data extracted from NIR, the pattern recognition methods including principal component analysis (PCA), linear Discriminant analysis (LDA) and support vector machine (SVM) were used. The results showed that Visible/Near Infrared (Vis/NIR) Spectroscopy was able to differentiate its lemon samples based on storage time. Although PCA, LDA and SVM methods were able to classify lemon samples with good accuracy according to qualitative characteristics, but LDA and SVM methods with 100% accuracy had better accuracy and fit. Also, according to the results, the quadratic function has been determined and introduced as the best function for constructing classification models by LDA and SVM methods.

ARTICLE INFO

Article History:

Received 2021/ 08/ 05
Accepted 2021/ 10/ 02

Keywords:

Near Infrared,
Visible Spectroscopy,
Lemon,
Storage Period.

DOI: 10.52547/fsct.18.120.26

DOR: 20.1001.1.20088787.1400.18.120.27.9

*Corresponding Author E-Mail:
movahhed@iauvaramin.ac.ir