

## سنجش کمی فنل کل انگور با استفاده از طیف‌سنجی فروسرخ نزدیک و شبکه عصبی مصنوعی

رضا محمدی گل<sup>۱\*</sup>، فرزاد آزادشهرکی<sup>۲</sup>، ولی اله لطفی<sup>۳</sup>

۱- استادیار، گروه مهندسی مکانیک بیوسیستم، دانشکده کشاورزی و منابع طبیعی، دانشگاه اراک، اراک، ایران  
\* نویسنده مسئول (r-mohammadigol@araku.ac.ir)

۲- استادیار، عضو هیأت علمی مؤسسه تحقیقات فنی و مهندسی کشاورزی، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران

۳- کارشناس ارشد، گروه مهندسی مکانیک بیوسیستم، دانشکده کشاورزی و منابع طبیعی، دانشگاه اراک، اراک، ایران

### چکیده

تاریخ دریافت: ۱۳۹۶/۰۱/۱۵

تاریخ پذیرش: ۱۳۹۶/۰۶/۰۸

### واژه‌های کلیدی

انگور

شبکه عصبی

طیف‌سنجی

غیرمخرب

فنل کل

انگور یکی از مهم‌ترین میوه‌ها در جهان است. ترکیبات فنلی، آنتی‌اکسیدان‌هایی هستند که از اجزاء مهم انگور بشمار می‌روند. اصطلاح ترکیبات فنلی شامل تمام مولکول‌های آروماتیکی از جمله اسیدهای آمینه تا مولکول‌های پیچیده شامل تانن‌ها و لگنین‌هاست. روش طیف‌سنجی فروسرخ نزدیک از رایج‌ترین روش‌های غیرمخرب سنجش ترکیبات و تعیین کیفیت میوه‌ها و سبزی‌هاست. در پژوهش حاضر امکان اندازه‌گیری فنل کل انگور توسط طیف‌سنجی فروسرخ نزدیک و شبکه عصبی مصنوعی (پرسپترون) مورد بررسی قرار گرفته است. تعداد ۴۴۴ نمونه (۱۰۷ نمونه رقم عسگری، ۱۰۶ نمونه رقم پیدانه قرمز، ۱۱۱ نمونه رقم شاه‌رودی و ۱۲۰ نمونه رقم خوش‌ناو) برای تدوین شبکه و اعتبارسنجی آن انتخاب شدند. شبکه‌های تدوین‌شده به‌وسیله شاخص انحراف پیش‌بینی باقی‌مانده در تخمین مقادیر فنل در نمونه‌های اعتبارسنجی (۱۰۱ نمونه) ارزیابی شدند. بیشترین مقدار برابر ۱/۶۶ در شبکه با توپولوژی ۱-۵-۸ با ضریب همبستگی (R) ۰/۷۹ و ریشه میانگین مربعات خطای پیش‌بینی ۴۸/۶۶ به‌دست آمد. نتایج نشان داد که امکان جداسازی مقدارهای کم از زیاد فنل کل با تکنیک طیف‌سنجی فروسرخ نزدیک و شبکه عصبی (پرسپترون) پس‌انتشار خطا به‌عنوان یک روش غیرمخرب وجود دارد.

### مقدمه

رنگ میوه‌ها و مقاومت آنها نسبت به بیماری دخالت دارند (احمدی و همکاران، ۱۳۸۸). ترکیبات فنلی اصلی انگور شامل کاتکین (کاتچین)<sup>۴</sup>، تانن<sup>۵</sup>، آنتوسیانین<sup>۶</sup>، فلاونل<sup>۷</sup>، هیدروکسی بنزوات<sup>۸</sup> و هیدروکسی سینامات<sup>۹</sup> هستند (Hemingway & Lakes, 1992).

ترکیبات فنلی آنتی‌اکسیدان‌هایی هستند که از اجزاء مهم انگور بشمار می‌روند. اصطلاح ترکیبات فنلی شامل تمام مولکول‌های آروماتیکی از جمله اسیدهای آمینه تا مولکول‌های پیچیده شامل تانن‌ها و لگنین‌ها هستند. تمام این ترکیبات از طریق مسیر ۳ مرحله‌ای آروماتیکی شیکیمات<sup>۱</sup>، فنیل پروپانوید<sup>۲</sup> و فلاونوئید<sup>۳</sup> ساخته می‌شوند. ترکیبات فنلی در طعم و مزه، گسی و

<sup>4</sup> Catechin

<sup>5</sup> Tanin

<sup>6</sup> Anthocyanin

<sup>7</sup> Flavonol

<sup>8</sup> Hydroxybenzoate

<sup>9</sup> Hydroxycinnamat

<sup>1</sup> Shikimate

<sup>2</sup> Phenylpropanoid

<sup>3</sup> Flavonoid

طیف‌سنجی NIR می‌تواند به‌عنوان یک روش سریع و کم‌هزینه برای تخمین ترکیبات فنلی در زمان تخمیر به‌کار گرفته شود.

Rolle و همکاران (۲۰۱۲) پیش‌بینی فنل هسته‌ای نوعی انگور<sup>۵</sup> را با استفاده از طیف‌سنجی FT-NIR (به کمک همبستگی و آنالیز واریانس) گزارش کردند. جمع‌بندی آنها بر این بوده است که طیف‌سنجی NIR می‌تواند به‌عنوان یک روش جایگزین، سریع، غیرمخرب و به‌صرفه برای پیش‌بینی مقدار فنل هسته‌ای انگور مورد تحقیق به‌کار گرفته شود.

همان‌طور که پیش‌تر اشاره شده واژه فنل کل مجموعه‌ای از آنتی‌اکسیدان‌ها بوده و وجود آنها ناظر بر کیفیت انگور است. تحقیق درباره کاهش زمان و هزینه اندازه‌گیری ترکیبات نسبت به روش‌های متداول (که جزء اهداف اصلی روش‌های اندازه‌گیری غیرمخرب هستند) در فرایند تولید محصول و صنایع تبدیلی مطالبه جهانی است. با عنایت به پیشرفت‌های صورت‌گرفته در عرصه فناوری و ساخت تجهیزات اپتیکی و نانو که باعث توسعه و ارتقاء چشمگیر طیف‌سنج‌های فرسوخ نزدیک شده است (تا آنجا که انواع قابل حمل آنها به بازار آمده است) تحقیق حاضر با هدف بررسی امکان سنجش کمی فنل کل انگور با تکنیک طیف‌سنجی فرسوخ نزدیک و استفاده از شبکه عصبی مصنوعی (پرسپترون<sup>۶</sup>) انجام شده است.

## مواد و روش‌ها

### تهیه نمونه

خوشه‌های انگور مورد نیاز از ایستگاه تحقیقات انگور وابسته به مرکز تحقیقات کشاورزی و منابع طبیعی استان قزوین تهیه شدند. تعداد ۴۸۰ خوشه از رقم عسگری، بیدانه قرمز، شاهرودی و خوشناو تهیه شد (هر رقم ۱۲۰ خوشه). به‌منظور ایجاد تنوع بهتر، خوشه‌ها از تاک‌های مختلف برداشت شدند. زمان برداشت برای هر رقم زمان برداشت مرسوم منطقه برای آن رقم بود. خوشه‌ها پس از برداشت به سردخانه با دمای ۱ درجه سلسیوس منتقل شدند. به‌منظور انتخاب نمونه از هر خوشه، ۱۰ حبه فاقد ترک، آسیب

ویژگی‌های کیفی و کمی میوه‌ها و سبزی‌ها را می‌توان توسط روش‌های مخرب و غیرمخرب اندازه‌گیری کرد. روش‌های غیرمخرب غالباً سریع بوده، ضایعات کمتری به همراه دارد و آسیبی به محصول نمی‌زنند (Hertog *et al.*, 2007). روش طیف‌سنجی فرسوخ نزدیک<sup>۱</sup> یکی از رایج‌ترین روش‌های غیرمخرب در تعیین کیفیت میوه‌ها و سبزی‌هاست و از آن در سنجش مقدار مواد جامد حل‌شدنی و pH ارقام مختلفی از میوه‌ها و سبزی‌ها استفاده شده است. از این روش در تعیین صفاتی که اندازه‌گیری آنها دشوار است، کمتر استفاده شده است (Nicolai *et al.*, 2007).

شبکه عصبی مصنوعی یکی از روش‌های مدل‌سازی غیرخطی هستند که در آن امکان مدل‌سازی یک تابع غیرخطی که بردار ورودی را به یک مقدار کمی نگاشت می‌کند، وجود دارد. به‌همین دلیل، شبکه عصبی به‌طور گسترده در سنجش‌های کیفی و کمی و تشخیص ترکیبات محصولات کشاورزی مورد استفاده قرار گرفته‌اند (محمدی‌گل و همکاران، ۱۳۹۴).

Ferrari و همکاران (۲۰۱۱) استفاده از طیف‌سنجی FT-NIR و HNMR را برای تفکیک نوشیدنی‌های<sup>۲</sup> انگور محتوی آنتوسیانین اصل (حاصل از انگور) و سطوح مختلف آنتوسیانین حاصل از برنج سیاه برای بهبود رنگ را گزارش کردند. مدل طبقه‌بندی آنها حد اقل مربعات جزئی<sup>۳</sup> بوده است. آنها گزارش کردند که کارایی NIR در تفکیک نمونه‌های تست ۷۰ درصد اما در مقابل طیف‌سنجی NMR در کنار انتخاب متغیرهای اصلی با تبدیل موجک<sup>۴</sup> (یکی از روش‌های کاهش داده‌ها به اطلاعات مفید و اثرگذار) بازده ۹۵ درصد به بالا داشته است.

Cozzolino و همکاران (۲۰۰۴) در تحقیق خود، طیف‌سنجی Vis-NIR به کمک روش مدل‌سازی PLS برای پیش‌بینی ترکیبات فنلی نوشیدنی انگور (۳ ترکیب اصلی در زمان تخمیر) را گزارش نمودند. آنها شاخص RPD در پیش‌بینی ۳ ترکیب را بین ۱/۸ و ۵/۸ گزارش نمودند و طبق جمع‌بندی آنها

<sup>1</sup> Near Infra Red (NIR)

<sup>2</sup> Wine

<sup>3</sup> Partial Least Square (PLS)

<sup>4</sup> Wavelet Transform

<sup>5</sup> Cabernet-Sauvignon

<sup>6</sup> Perceptron

براین اساس ۱۰ میکرولیتر از آب‌میوه با ۱۰۰ میکرولیتر معرف فولین سیوکالتیو و ۱۵۸۰ میکرولیتر آب‌مقطر مخلوط و به مدت ۵ دقیقه در شیکر قرار داده شد پس از آن ۳۰۰ میکرولیتر کربنات سدیم اشباع به آن اضافه گردید و پس از ۲ ساعت میزان جذب محلول در ۷۶۰ نانومتر در ۳ تکرار توسط میکروپلیت ریدر<sup>۳</sup> مدل Eon (BioTek, USA) قرائت گردید. میزان فنل کل عصاره براساس منحنی استاندارد و برحسب میلی‌گرم در لیتر غلظت اسیدگالیک<sup>۴</sup> بیان شد (Raja et al., 2014).

حذف داده‌های پرت<sup>۵</sup> و کاهش ابعاد داده‌های طیفی قبل از ایجاد ارتباط بین داده‌های طیفی و اندازه‌گیری‌های شیمیایی، داده‌های پرت حاصل از طیف‌برداری که ناشی از مشکلات فنی و طیف‌برداری نادرست بودند توسط روش آنالیز مؤلفه‌های اصلی<sup>۶</sup> شناسایی و حذف شدند. به این منظور نمونه‌های خارج از محدوده نرمال پراکندگی در نمودار امتیازدهی PCA به‌عنوان داده پرت شناخته شدند (Mireei et al., 2013). به‌طور کلی تعداد مشاهده‌ها (طیف‌ها) برای آموزش مؤفق یک مدل خودآموز ماشینی<sup>۷</sup>، با تعداد داده‌ها (متغیرهای طیفی) توسط رابطه‌نمایی افزایش می‌یابد (Ishikawa & Gulick, 2013). با توجه به حجم بالای متغیرهای طیف فرورسرخ نزدیک از مؤلفه‌های اصلی<sup>۸</sup> (PCs) به‌جای متغیرهای طیفی (مقدار بازتاب در هر طول موج) به‌عنوان ورودی شبکه استفاده شد (Mouazen et al., 2010).

محاسبه مؤلفه‌های اصلی از طیف‌های فرورسرخ نزدیک و به‌کارگیری آنها، در تحقیقات گوناگون مشاهده می‌شود (Mireei et al., 2013; Nicolai et al., 2007).

در نهایت ۴۴۴ نمونه (۱۰۷ نمونه رقم عسگری، ۱۰۶ نمونه رقم بیدانه قرمز، ۱۱۱ نمونه رقم شاه‌رودی و ۱۲۰ نمونه رقم خوشناو) برای تدوین مدل و

یا لهیدگی از یک قسمت خوشه (بالا یا پایین خوشه) انتخاب، جدا و به‌عنوان یک نمونه در نظر گرفته شد. حبه‌های هر خوشه به‌نحوی از خوشه جدا شدند تا کمترین آسیب به حبه‌ها رسیده و محل اتصال حبه‌ها به خوشه حفظ شد.

### طیف‌برداری و اندازه‌گیری فنل

پس از خروج خوشه‌ها از سردخانه و تهیه نمونه و قبل از برداشت طیف‌ها در هر روز، به مدت ۱۵ ساعت به نمونه‌ها فرصت داده شد تا دمای آنها به دمای محیط برسد. طیف‌سنج مورد استفاده در این پژوهش طیف‌سنج نوع فتودیودی (مدل EPP2000NIR (Stellarnet, USA) و مد مورد استفاده مد برهم‌کنش بوده است. در طیف‌سنج مورد استفاده آشکارساز از نوع ایندیم-گالیم-آرسناید<sup>۱</sup>، دامنه کاری ۹۰۰ تا ۱۷۰۰ نانومتر و وضوح طیف آشکارشده ۱۰۲۴ پیکسل بود. پیش از طیف‌سنجی، ابتدا طیف‌های تیره و سفید (مرجع) تعریف و ذخیره شدند. به این ترتیب که ابتدا با خاموش کردن منبع نور طیف تیره گرفته شد، سپس در حالت منبع نور روشن، از یک دیسک تفلون استاندارد که توانایی بازتاب ۹۵ درصد را در محدوده ۲۲۰۰-۲۵۰ نانومتر داشت برای دستیابی به طیف مرجع استفاده شد. بنابراین هنگام اخذ طیف از نمونه، نرم‌افزار طیف‌سنج میزان بازتاب نور را براساس رابطه (۱) محاسبه نمود:

رابطه (۱)

$$\%Tr_{\lambda} = \frac{S_{\lambda} - D_{\lambda}}{Ref_{\lambda} - D_{\lambda}} \times 100$$

در رابطه (۱)  $S_{\lambda}$  شدت نور خروجی حبه در طول موج  $\lambda$ ،  $Ref_{\lambda}$  شدت نور خروجی مرجع در طول موج  $\lambda$ ،  $D_{\lambda}$  شدت نور محیط در طول موج  $\lambda$  و  $Tr_{\lambda}$  میزان بازتاب نور در طول موج  $\lambda$  است.

پس از ثبت طیف از هر حبه در هر نمونه، میانگین طیف‌های حاصل از ۱۰ حبه یک نمونه به‌عنوان طیف آن نمونه در نظر گرفته شد. بعد از ثبت طیف از حبه‌ها بلافاصله نسبت به اندازه‌گیری فنل کل عصاره نمونه‌ها براساس روش فولین سیوکالتیو<sup>۲</sup> اقدام شد.

<sup>3</sup> Microplate reader

<sup>4</sup> Gallic acid

<sup>5</sup> Outliers

<sup>6</sup> Principle Component Analysis (PCA)

<sup>7</sup> Machine Learning Classifier

<sup>8</sup> Principal Components (PCs)

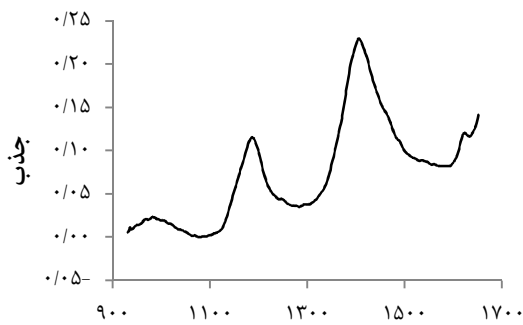
<sup>1</sup> Indium Gallium Arsenide

<sup>2</sup> Folin-Ciocaltiu

کاهش داده‌های طیفی به مؤلفه‌های اصلی (PCs) و تدوین شبکه‌ها در این تحقیق توسط نرم‌افزار IBM SPSS modeler نسخه ۱۴/۲ انجام پذیرفت.

### نتایج و بحث

در شکل (۱) میانگین طیف‌های جذبی نمونه‌ها در محدوده ۱۶۵۰-۹۳۰ نانومتر نشان داده شده است (طول موج‌های ابتدایی و انتهایی به دلیل وجود نویز حذف شدند). طیف جذبی دارای ۱ پیک بسیار پهن اطراف ۹۸۰ نانومتر وابسته به اورتون دوم OH، و پیک‌های مشخص دیگری وابسته به اورتون‌های دوم CH و اول OH به ترتیب در حدود ۱۲۰۰ و ۱۴۲۰ نانومتر و همچنین یک پیک کوچک در حدود ۱۶۳۰ نانومتر مربوط به اورتون اول CH مشاهده می‌شود (Cen & He, 2007).



شکل ۱- میانگین طیف‌های جذبی نمونه‌ها در محدوده ۱۶۵۰-۹۳۰ نانومتر

در جدول (۱) مشخصات آماری نمونه‌های به‌کاررفته در تحقیق پس از حذف داده‌های پرت با استفاده از تکنیک PCA آورده شده است. این روش توسط محققان دیگر هم استفاده شده است (Mouazen *et al.*, 2010). معمولاً ۲۵-۲۰ درصد داده‌ها به اعتبارسنجی مدل اختصاص داده می‌شوند (Nicolai *et al.*, 2007). در این پژوهش ۳۴۳ نمونه به تدوین مدل و ۱۰۱ نمونه به اعتبارسنجی آن اختصاص داده شد (جدول ۲).

اعتبارسنجی آن توسط نمونه‌های مستقل در نظر گرفته شدند.

### مدل‌سازی

برای تبدیل مفاهیم آماری طیفی به ارقام ریاضی مفهوم و استنباط ارتباط بین داده‌های طیفی و اندازه‌گیری‌های شیمیایی از مدل استفاده می‌شود (داده‌کاوی طیفی). در تحقیق حاضر از شبکه پرسپترون با یک‌لایه پنهان و قانون یادگیری پس‌انتشار خطا<sup>۱</sup> استفاده شد. شبکه‌های پس‌انتشار خطا رایج‌ترین و مناسب‌ترین شبکه‌ها برای کاربردهای طیف‌سنجی هستند (Mireei *et al.*, 2013). تعداد نرون‌های لایه پنهان توسط نرم‌افزار SPSS IBM modeler نسخه ۱۴/۲ با معیار کمترین خطای قابل دستیابی تعیین شده است.

### اعتبارسنجی

به‌منظور ارزیابی شبکه‌های تدوین‌شده از شاخص انحراف پیش‌بینی باقی‌مانده<sup>۲</sup> که از رابطه (۲) محاسبه می‌شود (Mouazen *et al.*, 2010; Jiang *et al.*, 2016) استفاده شد. از شاخص یادشده به‌عنوان بهترین معیار ارزیابی عملکرد مدل‌های رگرسیونی یادشده است (Mireei *et al.*, 2013).

رابطه (۲)

$$RPD = \frac{SD}{RMSE}$$

در رابطه (۲) SD انحراف معیار مقدارهای اندازه‌گیری شده و RMSE ریشه میانگین مربعات خطای پیش‌بینی<sup>۳</sup> است.

RMSE توسط رابطه (۳) محاسبه شد (Nicolai *et al.*, 2007).

رابطه (۳)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n_p} (\hat{y}_i - y_i)^2}{n_p}}$$

در رابطه (۳)  $y_i$  مقدار اندازه‌گیری‌شده فنل کل برای نمونه  $i$ ،  $\hat{y}_i$  مقدار پیش‌بینی‌شده صفت مورد نظر برای نمونه  $i$  و  $n_p$  تعداد نمونه است.

<sup>1</sup> Back propagation error

<sup>2</sup> Residual Prediction Deviation (RPD)

<sup>3</sup> Root Mean Square Error (RMSE)

جدول ۱- مشخصات آماری نمونه‌ها در اندازه‌گیری متغیر هدف

رقم	تعداد	فنل (میلی گرم در لیتر)		
		میانگین	انحراف معیار	مقدار بیشینه
عسگری	۱۰۷	۴۰۶/۱۰	۴۱/۹۶	۵۴۱/۲۴
بیدانه قرمز	۱۰۶	۳۷۴/۷۲	۴۱/۴۲	۴۷۱/۳۲
شاهرودی	۱۱۱	۴۰۳/۹۵	۶۶/۸۲	۶۷۱/۶۷
خوشناو	۱۲۰	۳۳۰/۷۴	۱۰۲/۵	۵۱۸/۰۱

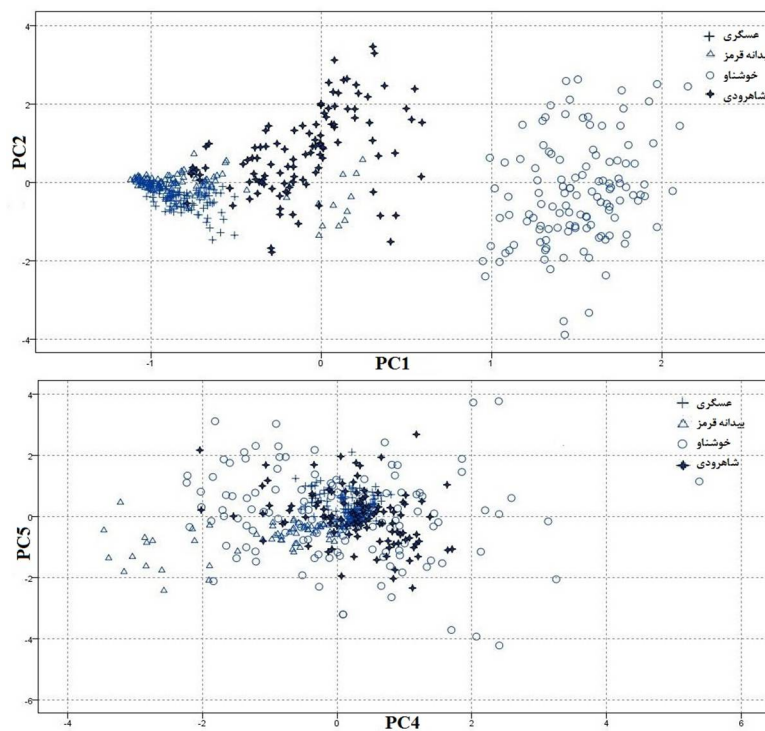
جدول ۲- مشخصات آماری مجموعه نمونه‌های مدل شده و نمونه‌های اعتبارسنجی مدل در اندازه‌گیری متغیر هدف

نمونه‌های اعتبارسنجی	تعداد	فنل (میلی گرم در لیتر)		
		میانگین	انحراف معیار	مقدار بیشینه
نمونه‌های مدل شده	۳۴۳	۳۷۶/۹۱	۷۴/۵۴	۶۷۱/۶۷
نمونه‌های اعتبارسنجی	۱۰۱	۳۸۰/۳۹	۸۰/۷۷	۵۴۱/۲۴

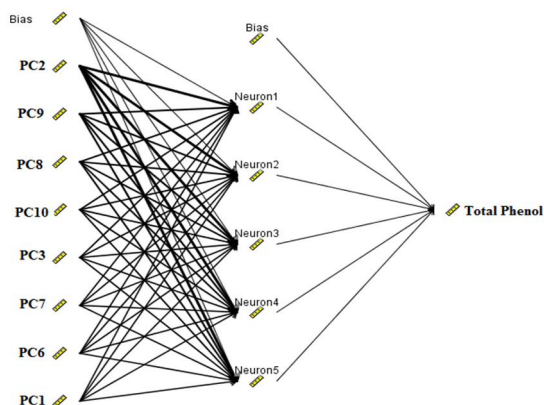
## جداسازی مؤلفه‌های اصلی کم اثر

یکی از تکنیک‌هایی که برای بالارفتن دقت مدل‌سازی مورد استفاده قرار می‌گیرد، به‌کارگیری مؤلفه‌های اصلی اثرگذار در فرایند تخمین متغیر هدف است. نرم‌افزار SPSS IBM Modeler نسخه ۱۴/۲ بر پایه الگوریتم‌های آماری درونی، متغیرهای مستقل را با درجه اهمیت آنها در تعیین فیلد هدف اولویت‌بندی می‌نماید. در این تحقیق، داده‌های طیفی به ۱۰ مؤلفه اصلی (از نظر اهمیت در تعیین کلاس طیف نمونه‌ها) کاهش یافته و پس از اولویت‌بندی در نرم‌افزار، ۸

مؤلفه اصلی به‌عنوان ورودی شبکه لحاظ شدند. این شیوه باعث بالارفتن دقت شده و توسط محققان دیگر نیز استفاده شده است (Feng et al., 2010). باتوجه به مطالب یادشده از ۱۰ مؤلفه اصلی، مؤلفه‌های اصلی ۴ و ۵ کنار گذاشته شدند. در شکل (۲) پراکنش طیف‌ها (قبل از مدل‌سازی) در مختصات مؤلفه‌های اصلی (۱ و ۲) و (۴ و ۵) مشاهده می‌شود. با در نظر گرفتن پراکنش نمونه‌ها، دلیل کنار گذاشتن مؤلفه‌های اصلی ۴ و ۵ از فرایند مدل‌سازی استنباط می‌شود.



شکل ۲- پراکنش نمونه‌ها (طیف‌ها) در مختصات مؤلفه‌های اصلی



شکل ۳- توپولوژی شبکه برتر در پیش‌بینی فنل از روی ۸ مؤلفه اصلی طیف NIR نمونه‌ها

### نتیجه‌گیری

نتایج تحقیق حاضر نشان داد شبکه عصبی پرسپترون را می‌توان در سنجش یا به تعبیر بهتر در تفکیک مقادیر کم و زیاد فنل کل انگور مورد استفاده قرار داد اما در اهداف سنجش کمی، دقت بالاتر نیاز بوده و RPD بیشتر از ۱/۸ لازم است (Jiang *et al.*, 2016; Mouazen *et al.*, 2010; Viscarra Rossel *et al.*, 2006). بنابراین تحقیق روی دیگر روش‌های مدل‌سازی داده‌های طیفی NIR یا افزایش و ارتقاء دقت شبکه عصبی مانند ترکیب شبکه با تکنیک‌های دیگر مدل‌سازی مانند الگوریتم ژنتیک یا استفاده از روش‌های مختلف پیش‌پردازش داده‌های طیفی به صورت منفرد و ترکیبی قبل از مدل‌سازی، پیشنهاد نویسندگان مقاله حاضر است.

حدود ۲۲ درصد از داده‌ها (معادل ۱۰۱ طیف) قبل از مدل‌سازی به طور کاملاً تصادفی از ۴ رقم برای اعتبارسنجی کنار گذاشته شدند (Han *et al.*, 2012). مشخصات آماری مجموعه نمونه‌های مدل‌سازی و نمونه‌های اعتبارسنجی شبکه در اندازه‌گیری متغیر هدف در جدول (۲) آمده است. باتوجه به اینکه در فرایند به‌کارگیری شبکه عصبی امکان اتفاقی بودن خروجی‌ها (نتایج) همیشه وجود دارد برای اطمینان از صحت نتایج به دست آمده، فرایند مدل‌سازی ۱۰ مرتبه انجام و سپس نتایج حاصله با مجموعه داده‌های اعتبارسنجی (که برای تمام شبکه‌ها یکسان بوده است) محاسبه شدند (جدول ۳).

نزدیک بودن ارقام در ستون‌های  $r$  و RPD جدول (۳) پدیده مطلوبی بشمار می‌آید. بیشترین RPD و  $r$  در توپولوژی شبکه ۱-۵-۸ (شکل ۳) با مقادیر به ترتیب ۱/۶۶ و ۰/۷۹ به دست آمده است. در تحقیق Muazen و همکاران (۲۰۱۰) به نقل از Viscarra Rossel و همکاران (۲۰۰۶) آمده است که چنانچه  $1/8 < RPD < 1/4$  باشد مدل نسبتاً خوب بوده و مدل را می‌توان در تفکیک مقادیر کم و زیاد متغیر هدف به کار برد. و اگر  $1/8 < RPD < 2$  به دست آید مدل خوب ارزیابی شده و سنجش کمی متغیر هدف امکان‌پذیر خواهد بود و چنانچه مقدار RPD بیشتر از ۲ محاسبه شود مدل نتایج دقیق‌تری در پی خواهد داشت.

جدول ۳- نتایج اعتبارسنجی شبکه‌های ایجاد شده

شبکه توپولوژی	$r$	RMSE	SD (Test samples)	RP D
۱-۷-۱	۰/۷۶	۵۲/۲۳	۸۰/۷۷	۱/۵۵
۱-۶-۱	۰/۷	۵۸/۱۶	۸۰/۷۷	۱/۳۹
۱-۵-۱	۰/۷۱	۵۶/۹۴	۸۰/۷۷	۱/۴۲
۱-۶-۱	۰/۷۱	۵۶/۵۵	۸۰/۷۷	۱/۴۳
۱-۵-۱	۰/۷۹	۴۸/۶۶	۸۰/۷۷	۱/۶۶
۱-۶-۱	۰/۷۵	۵۳/۷	۸۰/۷۷	۱/۵
۱-۴-۱	۰/۷۶	۵۲/۴	۸۰/۷۷	۱/۵۴
۱-۶-۱	۰/۷۷	۵۱	۸۰/۷۷	۱/۵۸
۱-۶-۱	۰/۷۲	۵۵/۴۶	۸۰/۷۷	۱/۴۶
۱-۶-۱	۰/۷۳	۵۶/۳۶	۸۰/۷۷	۱/۴۳

<sup>1</sup> Hold out

## منابع

- ۱- احمدی، ع.، احسان‌زاده، پ. و جباری، ف. ۱۳۸۸. مقدمه‌ای بر فیزیولوژی گیاهی. انتشارات دانشگاه تهران. ۶۵۳ ص.
- ۲- محمدی‌گل، ر.، خوش تقاضا، م.ه.، ملک فر، ر.، میرابوالفتحی، م. و نیکبخت، ع.م. ۱۳۹۴. تشخیص آفلاتوکسین پسته با استفاده از تکنیک طیف‌سنجی رامان و شبکه عصبی. نشریه ماشین‌های کشاورزی، ۵: ۹-۱.
- 3- Cen, H., & He, Y. 2007. Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food science and Technology*, 18(2): 72-83.
- 4- Cozzolino, D., Kwiatkowski, M.J., Parker, M., Cynkar, W.U., Damberg, R.G., Gishen, M., & Herderich, M.J. 2004. Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta*, 513(1): 73-80.
- 5- Feng, S., Chen, R., Lin, J., Pan, J., Chen, G., Li, Y., Cheng, M., Huang, Z., Chen, J., & Zeng, H. 2010. Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis. *Biosensors and Bioelectronics* 25(11): 2414-2419.
- 6- Ferrari, E., Focac, G., Vignalic, M., Tassia, L., & Ulrici, A. 2011. Adulteration of the anthocyanin content of red wines: Perspectives for authentication by Fourier Transform-Near InfraRed and 1H NMR spectroscopies. *Analytica Chimica Acta*, 701(2):139- 151.
- 7- Han, J., Kamber, M., & Pei, J. 2012. Data mining: concepts and techniques. Third edition „Morgan kaufmann Publishers (An imprint of Elsevier.), USA.
- 8- Hemingway, R.W., & Laks, P.E. (eds). 1992. Plant polyphenols synthesis, Properties, Significance. A division of plenum press, New York. N. Y. 1013-1053pp.
- 9- Hertog, M., Lammertyn, J., De Ketelaere, B., Scheerlinck, N., & Nicola, B.M. 2007. Managing quality variance in the postharvest food chain. *Trends in Food Science & Technology*, 18(6): 320-32.
- 10- Ishikawa, S., & Gulick, V. 2013. An automated mineral classifier using Raman spectra. *Computers & Geosciences*, 54: 259-268.
- 11- Jiang, Q., Chen, Y., Guo, L., Fei, T., & Qi, K. 2016. Estimating soil organic carbon of cropland soil at different levels of soil moisture using VIS-NIR spectroscopy. *Remote Sensing*, 8(9): 755-771.
- 12- Mireei, A., Mohtasebi, S.S., & Sadeghi, M. 2013. Comparison of linear and non-linear calibration models for non-destructive firmness determining of ‘Mazafati’ date fruit by NIR spectroscopy. *International Journal of Food Properties*. *International Journal of Food Properties*, 17(6): 1199-1210.
- 13- Mouazen, A.M., Kuang, B., De Baerdemaeker, J., & Ramon, H. 2010. Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma*, 158(1): 23-31.
- 14- Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Karen, I.T., & Lammertyn, J. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology*, 46(2): 99-118.
- 15- Raja, H.N., Dara, N.E., Hobaike, Z., Boussetta, N., Vorobiev, E., Maroun, R.G., & Louka, N. 2014. Extraction of total phenolic compounds, flavonoids, anthocyanins and tannins from grape byproducts by response surface methodology. Influence of solid-liquid ratio, particle size, time, temperature and solvent mixtures on the optimization process. *Journal of Food and Nutrition Sciences*, 5(04): 397-409.
- 16- Rolle, L., Torchio, F., Lorrain, B., Giacosa, S., Rio, S., Cagnasso, E., Gerbi, V., & Teissedre, P.L. 2012. Rapid methods for the evaluation of total phenol content and extract ability in intact grape seeds of cabernet-sauvignon: Instrumental mechanical properties and FT-NIR spectrum. *Journal International des Sciences de la Vigne et du Vin*, 46(1):29-40.
- 17- Viscarra Rossel, R.V., McGlynn, R.N., & McBratney, A.B. 2006. Determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy. *Geoderma*, 137(10):70-82.

## Quantification of Total Phenol in Grape by Near Infrared Spectroscopy and Artificial Neural Network

Reza Mohammadigol<sup>1\*</sup>, Farzad Azadshahraki<sup>2</sup>, Valiollah Lotfi<sup>3</sup>

1- Assistant Professor, Department of Biosystem Engineering, Faculty of Agriculture and Natural Resources, Arak University, Arak, Iran

\* Corresponding author (r-mohammadigol@araku.ac.ir)

2- Assistant Professor, Agricultural Engineering Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran

3- MSc, Department of Biosystem Engineering, Faculty of Agriculture and Natural Resources, Arak University, Arak, Iran

### Abstract

Grape is one of the most important fruits in the world. Phenolic compounds are antioxidants are important compositions of grape. Phenolic compounds phrase includes all the aromatic molecules consisting amino acids to complex molecules like tannins and lignin's. Near infrared spectroscopy is one of the most common nondestructive methods for fruits and vegetables qualification analysis. This research is conducted to evaluate the possibility of the quantification of total phenol in grape by near infrared spectroscopy and artificial neural network (perceptron). The number of 444 samples (107 Asgari, 106 Bidane, 111 shahroodi and 120 khoshnav varieties) were selected to model calibrating and test as well. Developed ANNs were compared on phenol prediction by residual prediction deviation (RPD) index in the test sample dataset (101 samples). The maximum RPD was 1.66 by 8-5-1 topology with correlation coefficient and root mean square (RMSE) equal to 0.79 and 48.66 respectively. It was concluded that NIR spectroscopy and back propagation perceptron ANN could be used to discriminate low and high amounts of grape total phenol as a nondestructive method.

**Keywords:** Grape, Neural Networks, Nondestructive, Spectroscopy, Total Phenol