پیش بینی گاف انرژی نیمه‌رساناهای بر پایه نیترید با استفاده از یادگیری ماشین

کاووس عباسی[۱]، انوشیروان غفاری‌پور[۲*]، جعفر جلیلیان[۱]

[۱] گروه فیزیک، دانشکده علوم پایه، دانشگاه یاسوج، یاسوج، ایران

[۲] گروه ریاضی، دانشکده علوم پایه، دانشگاه یاسوج، یاسوج، ایران

اطلاعات مقاله

تاریخ دریافت: ۱۴۰۳/۱۰/۱۶
تاریخ پذیرش: ۱۴۰۳/۱۱/۱۷
تاریخ چاپ: ۱۴۰۳/۱۱/۲۰

شاپای چاپی: 2588-493x
شاپای الکترونیکی: 2588-4921

* نویسنده مسئول
ahgaffaripour@yu.ac.ir

چکیده

در این مقاله با استفاده از شبیه سازی کوانتومی مبتنی بر نظریه تابعی چگالی، گاف انرژی ۳۰۰ ترکیب نیتریدی مورد بحث و بررسی قرار گرفت. گاف انرژی ترکیبات با استفاده از دو تقریب GGA-PBE و HSE06 محاسبه شده است. پارامترهای مورد بررسی در مطالعات یادگیری ماشین به دو دسته پارامترهای اتمی و کریستالی دسته بندی شده‌اند. پارامترهای اتمی عبارتند از شعاع کووالانسی، الکترونگاتیوی، تعداد الکترون‌های ظرفیت و اولین انرژی یونیزاسیون. بعد از جمع آوری داده‌های ویژگی‌های اتمی مدل رگرسیون خطی چندگانه به داده ها برازش داده شد. در ادامه با روش رگرسیون گام به گام با معیار AICc انتخاب متغیر انجام شده است و اندازه اثر ویژگی‌های مختلف محاسبه شده است. همچنین برای افزایش دقت مدل در پیش بینی HSE، سه ویژگی بلوری نیز در مدل گنجانده شده اند و ۸ مدل مختلف هر کدام با حضور یک یا چند مورد از سه ویژگی بلوری فوق برازش داده شده است. یافته ها نشان می‌دهد مدلی که هیچکدام از متغیرهای بلوری در آن حضور ندارند دارای ضریب تعیین تعدیل شده (R^2) به میزان به ترتیب 75.45% است که با اضافه شده ویژگی های بلوری، نتایج به میزان قابل توجهی بهبود پیدا می کند. به طور مشخص اضافه شدن متغیر گاف انرژی PBE باعث افزایش R^2 به میزان 99.03% (از 75.45% به 99.03%) دارد.

واژگان کلیدی: گاف انرژی، نظریه تابعی چگالی، نیمه‌رسانا، یادگیری ماشین، رگرسیون خطی

مقدمه

پیش‌بینی گاف انرژی نیمه‌رساناها با استفاده از یادگیری ماشین یکی از حوزه‌های مهم و نوظهور در علوم مواد و فیزیک حالت جامد است. نیمه‌رساناها به دلیل کاربردهای گسترده‌شان در الکترونیک و اپتوالکترونیک همواره مورد توجه پژوهشگران قرار گرفته‌اند. گاف انرژی، که به اختلاف انرژی بین نوار ظرفیت و نوار رسانش اشاره دارد، یکی از ویژگی‌های بنیادی این مواد است و نقش تعیین‌کننده‌ای در رفتار الکتریکی و اپتیکی آنها ایفا می‌کند. پیش‌بینی دقیق این ویژگی، به ویژه برای مواد جدید و کمتر مطالعه‌شده، اهمیت زیادی در طراحی و مهندسی مواد پیشرفته دارد. روش‌های رایج برای محاسبه گاف انرژی معمولاً مبتنی بر شبیه‌سازی‌های دقیق کوانتومی نظیر نظریه تابعی چگالی هستند که با وجود دقت بالا، نیازمند زمان و منابع محاسباتی زیادی هستند. این محدودیت‌ها به‌ویژه در بررسی

[Downloaded from jmrph.khu.ac.ir on 2026-07-01]

[DOI: 10.61882/jmrph.9.1.27]

تعداد زیادی ترکیب شیمیایی یا ساختارهای پیچیده، مشکل‌ساز می‌شوند. در این میان، یادگیری ماشین به عنوان رویکردی قدرتمند و کارآمد مطرح شده است که می‌تواند از داده‌های موجود برای ایجاد مدل‌هایی استفاده کند که با دقت و سرعت بالا گاف انرژی را پیش‌بینی کنند.

مدل‌های یادگیری ماشین، نظیر رگرسیون، شبکه‌های عصبی مصنوعی، و روش‌های مبتنی بر درخت تصمیم، با استفاده از داده‌های تجربی و شبیه‌سازی‌شده آموزش می‌بینند. این مدل‌ها قادرند الگوهای پیچیده‌ای را که ممکن است در روش‌های تحلیلی سنتی ناشناخته بمانند، شناسایی کنند. علاوه بر این، استفاده از یادگیری ماشین امکان کشف مواد جدید با ویژگی‌های مطلوب را از طریق غربالگری داده‌ها فراهم می‌کند. به عنوان نمونه، مطالعه‌ای توسط Ruppو همکارانش در سال ۲۰۱۲ نشان داده است که مدل‌های یادگیری ماشین می‌توانند با دقت بالا خواص مولکولی، از جمله گاف انرژی ترکیبات را پیش‌بینی کنند [۱]. همچنین، در پژوهشی دیگر، Ren و همکاران با ترکیب یادگیری ماشین و آزمایش‌های با توان بالا توانسته‌اند کشف مواد جدید را بر اساس داده های موجود پیش بینی کنند و تسریع ببخشند[2] .

پیشرفت‌های اخیر در این زمینه نشان‌دهنده پتانسیل بالای این رویکرد در بهبود سرعت و دقت پیش‌بینی‌ها است. با این حال، چالش‌هایی نظیر کیفیت و کمیت داده‌های آموزشی، انتخاب ویژگی‌های مؤثر، و تفسیرپذیری مدل‌ها همچنان وجود دارند. درک و رفع این چالش‌ها می‌تواند مسیر را برای استفاده گسترده و کاربردی از یادگیری ماشین در پیش‌بینی گاف انرژی و سایر ویژگی‌های مواد هموار کند.

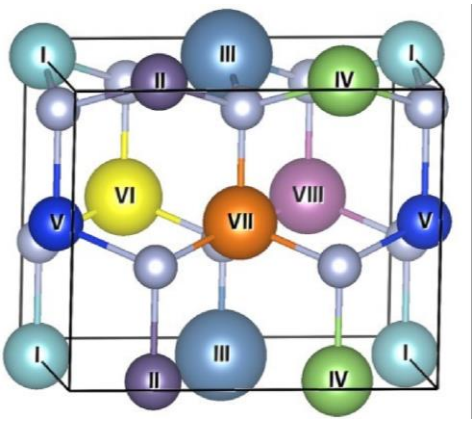
۱- روش تحقیق

در این پژوهش ساختار ابرياخته ۱۶ اتمی 2×2×2 ورتسایت GaN مورد بررسی و مطالعه قرار گرفته است. بطوریکه کاتیون های مختلفی امکان جایگزینی در ساختار مورد نظر به جای اتم Ga را دارا می باشند. این کاتیون ها شامل اتم های گروه IIA نظیر Be^{2+}، Mg^{2+}، Ca^{2+}، Sr^{2+} و Ba^{2+} و گروه IIB شامل Zn^{2+} و Cd^{2+}هستند. کاتیون‌های 3+ از عناصر گروه IIIA شامل Al^{3+}، Ga^{3+} و In^{3+} و گروه IIIB شامل Sc^{3+} و Y^{3+} تشکیل شده‌اند. همچنین، کاتیون‌های 4+ از عناصر گروه IVA شامل Si^{4+}، Ge^{4+} وSn^{4+} و گروه IVB شامل Ti^{4+}، Zr^{4+} و Hf^{4+} انتخاب شده‌اند.

نحوه جایگذاری این کاتیون های در ساختار ابریاخته مورد نظر به صورت زیر می باشند. جای‌گذاری کاتیون‌ها به سه نوع تقسیم شده است که در شکل ۱ ارائه شده است. اصول این تقسیم‌بندی به ترتیب زیر است. نوع اول: موقعیت‌های II، III، V و VII توسط کاتیون‌های 2+ و موقعیت‌های II، IV، VI و VIII توسط کاتیون‌های 4+ اشغال می‌شوند. نوع دوم :هر هشت موقعیت کاتیونی تنها توسط کاتیون‌های 3+ اشغال می‌شوند .نوع سوم: موقعیت‌های I و III توسط کاتیون‌های 2+، موقعیت‌های II وIV توسط کاتیون‌های 4+ و سایر موقعیت‌ها توسط کاتیون‌های 3+ اشغال می‌شوند. در مجموع، 68115 ترکیب مختلف از نیتریدها ایجاد شدند. از این میان، 300 ترکیب به صورت تصادفی انتخاب شده [3] و گاف آن‌ها با استفاده از روش هیبریدی HSE06 محاسبه شده است. برای ترکیبات نیترید طراحی‌شده در این مطالعه، به دلیل الکترونگاتیوی بالای اتم نیتروژن، انرژی پیوند سیستم معمولاً بالا است و انتظار می‌رود که ساختارهای نیتریدی در فضای طراحی از نظر ساختاری پایدار باشند. بر اساس محاسبات شبیه سازی انجام شده انرژی بستگی ۳۰۰ ماده انتخاب‌شده به صورت تصادفی همگی منفی هستند. تمامی نتایج محاسباتی در جدول 1 در اطلاعات تکمیلی فهرست شده‌اند.

[Downloaded from jmrph.khu.ac.ir on 2026-07-01]

[DOI: 10.61882/jmrph.9.1.27]



شکل ۱: ساختار کریستالی ترکیب در نظر گرفته شده و جایگاه های تفاوت اتمی برای آلاییدن یون‌های مورد نظر.

محاسبات اصول اولیه با استفاده از نظریه تابعی چگالی (DFT) و نرم‌افزار شبیه‌سازی **کوانتوم اسپرسو** [4] انجام گرفت است.. توابع موج الکترون‌های ظرفیت در چارچوب مدل موج تخت بسط داده شده و برای اطمینان از دقت بالای محاسبات، انرژی قطع بردار موج الکترونی و بردار بسط چگالی به ترتیب 60 و 300 ریدبرگ در نظر گرفته شده است. برای بهبود دقت محاسبات تبادلی-همبستگی، از تابع گرادیان تعمیم‌یافته (GGA) که توسط Perdew, Burke, and Ernzerhof (PBE) ارائه شده است و همچنین تقریب تابع هیبریدی HSE06 استفاده گردید [4] همچنین برای بهینه‌سازی ساختارها، تمام سیستم‌ها تا زمانی که نیروهای وارد بر هر اتم به کمتر از ۰،۰۱ الکترون‌ولت بر آنگستروم برسد، به طور کامل بهینه‌سازی شدند. برای مش بندی فضای وارون، از روش **مون‌خورست-پک** با ابعاد 8 × 6 × 8 استفاده شده است.

از تابع هیبریدی HSE06 برای محاسبه دقیق گاف انرژی استفاده شده است. برای محاسبه اختلاف انرژی (band offset) ، 300 ابریاخته از ساختارهای نیتریدی در جهت (001) ساخته شدند. این ابریاخته به شکل $(XN)_n/(GaN)_n$ است، که در آن XN نمایانگر هر یک از ۳۰۰ نیتریدی است که به صورت تصادفی انتخاب شده‌اند و n برابر با 5 است. یعنی در هر ابریاخته،5 لایه نیترید و گالیم نیترید به طور متناوب استفاده شده است. در نهایت، گاف انرژی برای این ساختارها با استفاده از تقریب HSE06 و PBE محاسبه شدند.

۲- بحث و نتیجه گیری

انرژی تشکیل E_f برای هر ترکیب نیتریدی با استفاده از معادله (۱) به دست آمد. این معادله انرژی ترکیب نیترید را از انرژی عناصر اولیه‌اش محاسبه می‌کند و به طور کلی برای بررسی پایداری ترکیب‌ها و مقایسه با دیگر ترکیبات به کار می‌رود. که در این معادله E_c انرژی کل ترکیب و E_i انرژی مربوط به عناصر سازنده آن ترکیب است.

$$E_f = E_C - \sum_{I}^{VIII} E_i \qquad (1)$$

در واقع ΔE_V اختلاف نوارهای انرژی بین دو ماده نیمه‌رسانا را نشان می دهد که به وسیله معادله (۲) محاسبه می شود.

$$\Delta E_V(XN/GaN) = \Delta E_{V,C'}(XN) - \Delta E_{C'/C}(XN/GaN) + A_V(XN) + A_V(GaN) \qquad (2)$$

$$\Delta E_{V,C'}(XN) = E_V(XN) - E_{C'}(XN) \qquad (3)$$

$$\Delta E_{V,C}(GaN) = E_V(GaN) - E_C(GaN) \qquad (4)$$

E_V انرژی مربوط به نوار ظرفیت در مواد مجزا است، E_C $(E_{C'})$ سطح انرژی مغزه ها در مواد مجزا را نشان می‌دهد . $\Delta E_{C'/C}$ اختلاف سطح انرژی مغزه ها بین دو ماده در دو طرف یک مدل فصل مشترک ساخته‌شده را توصیف می‌کند، و A_V پتانسیل تغییرشکل نوارظرفیت است. سطح تراز انرژی مغزه ها به‌عنوان سطح اوربیتال 1S اتم نیتروژن تنظیم شده است، به‌دلیل انرژی نسبتاً پایین آن که حدود ۳۷۰- الکترون‌ولت است. با در نظر گرفتن شباهت‌های هندسی بین نیتریدهای ساخته‌شده در فضای ساختاری، برای ساده‌سازی فرض شد که پتانسیل تغییرشکل نوارظرفیت هر ترکیب زمانی که با GaN ورتسایت متصل می‌شود، نادیده گرفته می‌شود؛ یعنی

[Downloaded from jmrph.khu.ac.ir on 2026-07-01]

[DOI: 10.61882/jmrph.9.1.27]

$A_V(GaN)+A_V(XN)= A_V(GaN) + A_V(XN) = 0$ در نتیجه، اختلاف نوار بین XN ورتسایت و GaN به‌صورت زیر نوشته می‌شود:

$E_{V,C'}(XN)-\Delta E_{V,C}(GaN)+\Delta E_{C'/C}(XN/GaN).$ اختلاف نوارهای محاسبه‌شده با نتایج گزارش‌شده است.

انتخاب پارامترهای ورودی برای محاسبات یادگیری ماشین بر اساس ویژگی ها تک عنصری و ویژگی های بلوری ترکیبات انجام شده است. از ویژگی‌های تک اتمی، شعاع کووالانسی R، الکترونگاتیوی X، تعداد الکترون‌های ظرفیت V و انرژی یونش اول I، برای تمامی اتم‌های موجود در هر ترکیب مورد نظر قرار گرفته است. همچنین پارامترهای کریستالی انرژی تشکیل، لبه نواری انرژی و همچنین گاف انرژی حاصل از محاسبات PBE را در گام دوم به عنوان پارامترهای ورودی یادگیری ماشین وارد محاسبات کرده‌ایم.

دلیل انتخاب پارامترهای کریستالی مورد مطالعه در این پژوهش به شرح زیر می باشد. انرژی تشکیل بلور معیاری از پایداری ترکیبات می باشد و هر چقدر این کمیت بزرگ تر بوده، به معنی پایداری کریستالی بیشتر ترکیب می باشد. انتخاب این پارامتر به این خاطر بوده است که از لحاظ پایداری هم نقش این ویژگی بلوری مهم در محاسبات یادگیری ماشین بررسی شود. کمیت دیگر انرژی لبه نوار ترکیبات است، که کمیتی مهم و وابسته به گاف انرژی ترکیبات و همچنین عناصر تشکیل دهنده آنها دارد. این کمیت به طریقی به مقدار الکترونگاتیوی تمامی عناصر تشکیل دهنده یک ترکیب وابسته است. به همین خاطر وارد کرده آن علاوه بر مرتبط بودن با مقدار پارامتر الکترونگاتیوی، بیانگر رابطه الکترون خواهی و الکترون دهی عناصر مختلف در ترکیبات می باشد.

در نهایت کمیت گاف انرژی ترکیبات با استفاده از تقریب PBE را به عنوان ویژگی بلوری در محاسبات یادگیری ماشین وارد کرده این. هدف اصلی این پژوهش پیش بینی گاف انرژی ترکیبات بصورت دقیق می باشد که در اینجا مبنای و ملاک سنجش ما، تقریب HSE06 می باشد. هر دوی تقریب های PBE و HSE06 گاف انرژی را پیش بینی می کنند. اما با اختلاف بسیار زیاد. هدف از وارد کردن این پارامتر این بود که محاسبات تقریب PBE برای سنگین ترین ترکیبات هم در زمان بسیار کمی قابل محاسبه می باشد. در حالیکه برای HSE06 می تواند روزها به طول بینجامد. به همین خاطر استفاده از نتایج PBE در روند یادگیری مشاین می تواند کمک شایان توجهی به پیش بینی دقیق تر نتایج یادگیری ماشین شود.

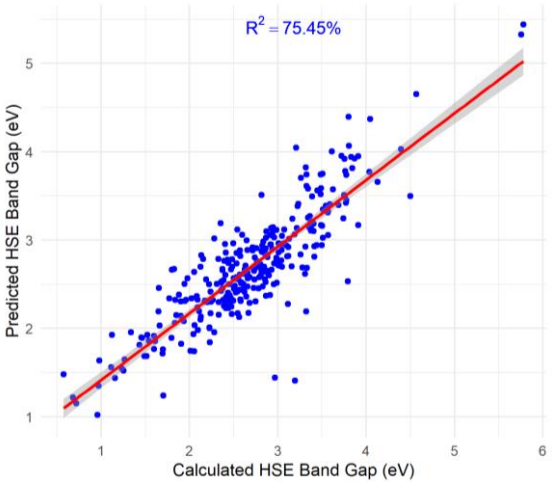
نتایج حاصل از محاسبات یادگیری ماشین بر روی داده های محاسباتی انجام شده را در ادامه می بینیم. بعد از جمع آوری داده های ویژگی‌های اتمی مدل رگرسیون خطی چندگانه به داده ها برازش داده شد. در ادامه با روش رگرسیون گام به گام با معیار AICc انتخاب متغیر انجام شده است و اندازه اثر ویژگی های مختلف محاسبه شده است. همچنین برای افزایش دقت مدل در پیش بینی HSE06، سه ویژگی بلوری نیز در مدل گنجانده شده اند و ۸ مدل مختلف هر کدام با حضور یک یا چند مورد از سه ویژگی فوق برازش داده شده است. یافته ها نشان می دهد مدلی که هیچکدام از متغیرهای بلوری در آن حضور ندارند دارای ضریب تعیین تعدیل شده (R^2) به میزان 75.45% است که با اضافه شده ویژگی های بلوری، نتایج به میزان قابل توجهی بهبود پیدا می‌کند. به طور مشخص اضافه شدن متغیر PBE باعث افزایش R^2 به میزان 23.58% (از 74.45% به 99.03%) دارد.

با نگاهی به شکل‌های ۲ تا ۹ حاصل از رگرسیون خطی داده های شبیه سازی برای پیش بینی گاف انرژی، می بینیم که تقریبا انرژی تشکیل و انرژی لبه نواری در ترکیبات مورد بررسی به عنوان داده های کریستالی ورودی برای پیش بین گاف انرژی تاثیر بسیار ناچیزی دارند. از طرف دیگر زمانی که گاف انرژی حاصل از نتایج PBE به داده های ورودی افزوده می شود، درصد تطابق داده های پیش بینی با نتایج محاسباتی به 99.03% می رسد. از آنجا که محاسبات GGA-PBE

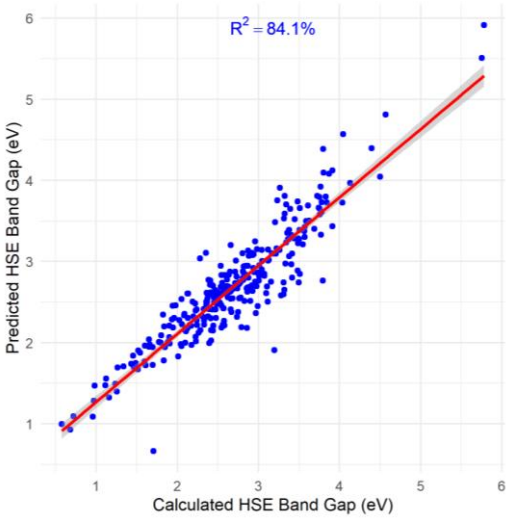
[Downloaded from jmrph.khu.ac.ir on 2026-07-01]

[DOI: 10.61882/jmrph.9.1.27]

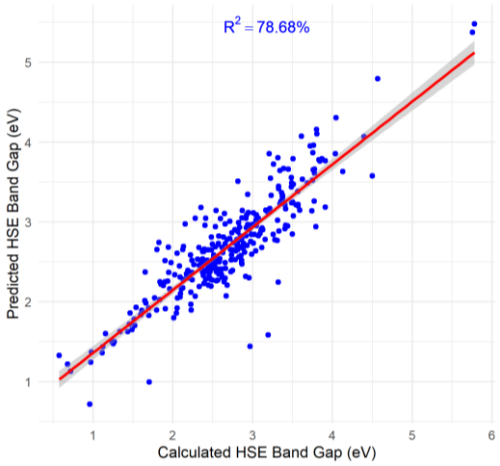
اط لحاظ زمانی بسیار سریع تر از محاسبات HSE06 صورت می پذیرد، لذا می توان با داشتن پارامترهای اتمی ترکیبات و مقدار گاف حاصل از تقریب GGA، مقدار گاف دقیق ترکیبات را با استفاده از یادگیری ماشین پیش بینی نمود.



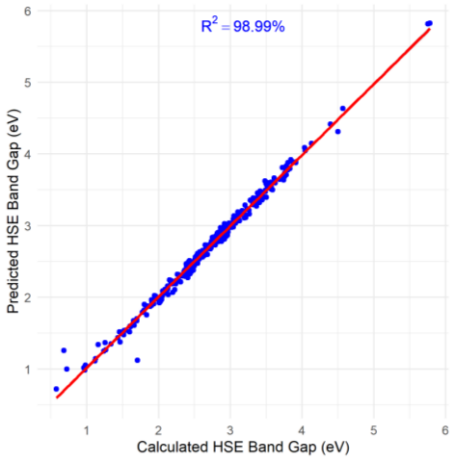
شکل **2**: نتایج حاصل برای یادگیری ماشین با درنظر گرفتن تمامی پارامترهای اتمی



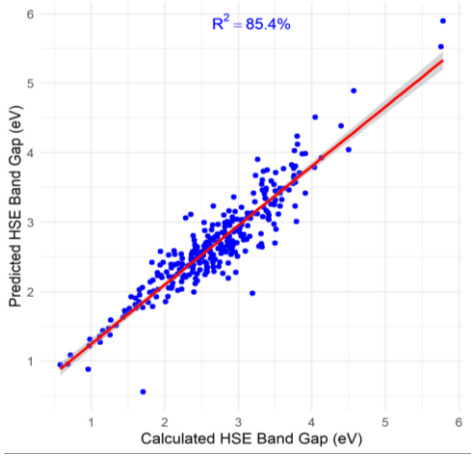
شکل ۳: نتایج حاصل برای یادگیری ماشین با درنظر گرفتن تمامی پارامترهای اتمی به علاوه انرژی نواری



شکل **4**: نتایج حاصل برای یادگیری ماشین درنظر گرفتن تمامی پارامترهای اتمی به علاوه انرژی تشکیل



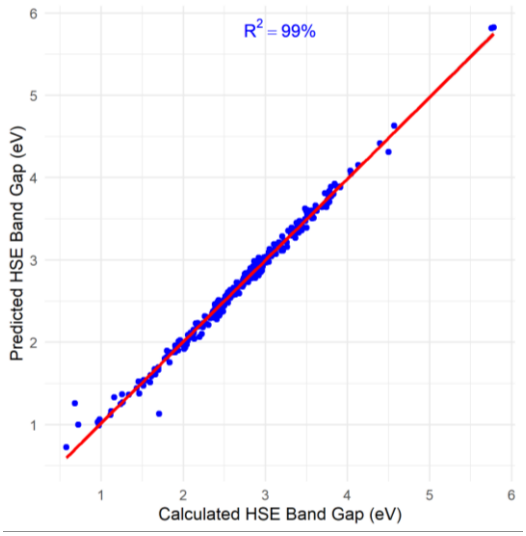
شکل ۵: نتایج حاصل برای یادگیری ماشین درنظر گرفتن تمامی پارامترهای اتمی به علاوه گاف انرژی **PBE**



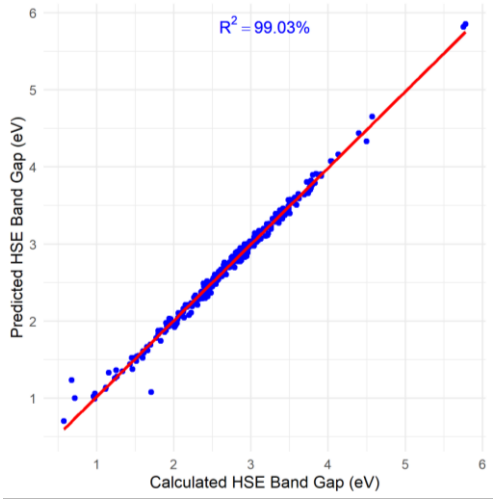
شکل ۶: نتایج حاصل برای یادگیری ماشین درنظر گرفتن تمامی پارامترهای اتمی به علاوه انرژی نواری و انرژی تشکیل

[Downloaded from jmrph.khu.ac.ir on 2026-07-01]

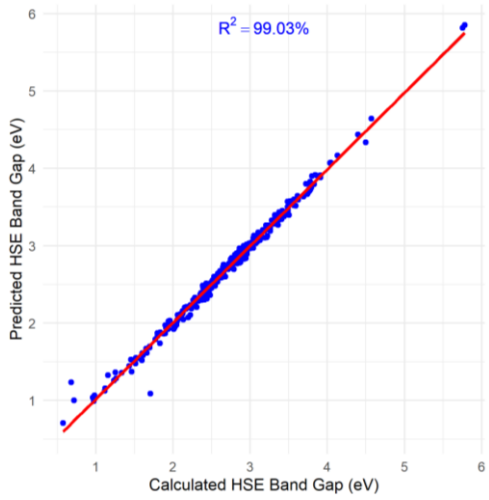
[DOI: 10.61882/jmrph.9.1.27]



شکل ۷: نتایج حاصل برای یادگیری ماشین درنظر گرفتن تمامی پارامترهای اتمی به علاوه گاف انرژی و انرژی تشکیل



شکل ۸: نتایج حاصل برای یادگیری ماشین درنظر گرفتن تمامی پارامترهای اتمی به علاوه گاف انرژی و انرژی نواری



شکل ۹: نتایج حاصل برای یادگیری ماشین درنظر گرفتن تمامی پارامترهای اتمی و کریستالی

۳- نتیجه گیری

با استفاده از روش رگرسیون خطی و استفاده از داده های شبیه سازی کوانتومی مبتنی بر نظریه تابعی چگالی، گاف انرژی ترکیبات پایه نیتریدی پیش بینی شد. گاف انرژی حاصل از نتایج محاسبات HSE06 به عناون داده مرجع انتخاب شدند. دو دسته داده اولیه ورودی، داده های اتمی و داده های کریستالی انتخاب شدند. نتایج حاصل از یادگیری ماشین نشان داد که پیش بینی گاف انرژی ترکیبات با استفاده از داده های اتمی حدود ۷۵ درصد تطابق با نتایج حاصل از شبیه سازی کوانتومی داشت. ایین در حالیست که با افزودن مقدار گاف انرژی حاصل از نتایج GGA-PBE، این درصد به مقدار ٪۹۹ درصد رسیده. دلیل این افزایش دقت نتایج یادگیری ماشین با حضور مقدار پارامتر گاف انرژی **PBE** می تواند این باشد که هر دوی این کمیت های یعنی گاف انرژی **PBE** به عنوان ورودی و کمیت گاف انرژی **HSE06** به عنوان کمیت پیش بینی، کمیت فیزیکی مشابهی را پیش بینی می کنند. به همین خاطر است که

[Downloaded from jmrph.khu.ac.ir on 2026-07-01]

[DOI: 10.61882/jmrph.9.1.27]

وارد کردن مقدار **PBE** نتایج را با دقت ۹۹ درصد پیش بینی می کند.

نتایج حاصل از این تحقیق نشان می دهد که برای ترکیبات نیمه رسانای نیتریدی می توان با داشتن گاف انرژی تقریب GGA-PBE، مقدار گاف انرژی دقیق تقریب HSE06 را پیش بینی نمود.

منابع

[1] K. Hansen et al., "Machine learni predictions of molecular properti Accurate many-body potentials a nonlocality in chemical space," *J. Ph Chem. Lett*. vol. 6 2326−2331, 2015.

[۲] F. Ren et al., "Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments," *Sci. Adv.* vol. 4, 1566, 2012.

[۳] Y. Huang et al., "Band gap and band alignment prediction of nitride based semiconductors using machine learning", *J. Mater. Chem. C*, vol. **7,** 3238-3245, 2019.

[Downloaded from jmrph.khu.ac.ir on 2026-07-01]

[DOI: 10.61882/jmrph.9.1.27]

JMRPh. (9)(1): 27-34, Spring and Summer, 2024

Prediction of the Energy Gap of Nitride-Based Semiconductors Using Machine Learning

Article details

Received: 2025/01/5
Accepted: 2025/02/5
Published: 2025/02/9

ISSN: 2588-493x
eSSN: 2588-4821

Correspondence email:
ahgaffaripour@yu.ac.ir

Abstract

In this article, the energy gap of 300 nitride compounds was analyzed and investigated using quantum simulations based on Density Functional Theory (DFT). The energy gaps of the compounds were calculated using two approximations, GGA-PBE and HSE06. The parameters considered in the machine learning studies were categorized into two groups: atomic and crystalline parameters. The atomic parameters include covalent radius, electronegativity, the number of valence electrons, and the first ionization energy. After collecting data on atomic features, a multiple linear regression model was fitted to the data. Subsequently, variable selection was performed using the stepwise regression method with the AICc criterion, and the effect size of various features was calculated. Additionally, to improve the accuracy of the model in predicting HSE06, three crystal features were incorporated into the model, and eight different models were fitted, each including one or more of these crystal features. The findings indicate that the model without any crystal variables has an adjusted coefficient of determination (R^2) of 75.45%. However, with the inclusion of crystal features, the results improve significantly. Specifically, adding the PBE energy gap as a variable increases the R^2 to 99.03% (from 75.45% to 99.03%).

Keywords: Energy Gap, Density Functional Theory, Semiconductor, Machine Learning, Linear Regression.

[Downloaded from jmrph.khu.ac.ir on 2026-07-01]

[DOI: 10.61882/jmrph.9.1.27]