

استفاده از روش ترکیبی موجک-ماشین بردار پشتیبان برای پیش‌بینی وقوع امواج غیرعادی

کیومرث محمودی¹، محمد جواد کتابداری^{2*} و حسن قاسمی³

1- دانشجوی دکتری، دانشکده مهندسی دریا، دانشگاه صنعتی امیرکبیر، تهران

2- دانشیار، دانشکده مهندسی دریا، دانشگاه صنعتی امیرکبیر، تهران

3- استاد، دانشکده مهندسی دریا، دانشگاه صنعتی امیرکبیر، تهران

*ketabdar@aut.ac.ir

چکیده- امواج غیرعادی یا سهمگین به امواجی گفته می‌شود که با ارتفاع خیلی زیاد، ولی بطور ناگهانی و غیر قابل پیش‌بینی و به ندرت رخ می‌دهند. عوامل مختلفی مثل طوفان‌های شدید، توپوگرافی خاص بستر، تلاقی جریان‌های کرانه‌ای و امواج و برهم‌کنش امواج با طول موج‌ها و فرکانس‌های مختلف با یکدیگر، ممکن است سبب بروز آن‌ها شوند. اما همه اینها هنوز در حد فرضیه هستند. هدف از این تحقیق ارائه یک روش ترکیبی برای پیش‌بینی وقوع امواج غیرعادی از روی متغیرهای هواشناسی با استفاده از روش ترکیبی تبدیل موجک و ماشین بردار پشتیبان دسته‌بندی کننده است. برای بررسی کارایی مدل ارائه شده از داده‌های تاریخی متغیرهای هواشناسی برداشت شده طی سه سال از ایستگاه‌های شماره 41041 و 41004 در دو طوفان مشهور Dean (2007) و Irene (2011) استفاده شده است. ابتدا با استفاده از تبدیل موجک امواج غیرعادی شناسایی شده، سپس از خروجی‌های این روش برای آموزش ماشین بردار پشتیبان استفاده شده است. تبدیل موجک روی داده‌های ارتفاع مشخصه اعمال می‌شود. در این روش امواج غیرعادی نواحی از طیف تبدیل موجک را تشکیل می‌دهند که در آن‌ها انرژی زیاد است. از داده‌های تاریخی هواشناسی پارامترهایی مانند جهت باد، سرعت باد، ارتفاع موج مشخصه، فشار سطح دریا، دمای هوا و دمای سطح دریا که در طول سه سال بصورت ساعتی ثبت شده، برای آموزش و آزمون مدل‌های ماشین بردار پشتیبان استفاده شده است. نتایج نشان داد که روش ارائه شده از دقت قابل قبولی برای پیش‌بینی وقوع امواج غیرعادی برخوردار است.

کلید واژگان: پیش‌بینی امواج غیرعادی، داده‌های هواشناسی، ماشین بردار پشتیبان، تبدیل موجک.

1- مقدمه

تلاش‌های زیادی برای شناسایی امواج دریاها صورت گرفته است، اما ماهیت موج آن‌گونه که باید، شناخته نشده است. این مسئله زمانی که نیروهای موثر افزایش می‌یابد، بسیار پیچیده‌تر شده و سبب وقوع امواج غیرعادی یا سهمگین¹ می‌شود. شناسایی و پیش‌بینی شرایط وقوع امواج غیرعادی بسیار دشوارتر از سایر امواج است. به همین دلیل در تحقیقات بسیار کمتر به آن توجه شده است. پیش‌بینی شرایط وقوع امواج غیرعادی بسیار حائز اهمیت

سطح دریاها و اقیانوس‌ها همواره پوشیده از امواج با بسامدها و طول موج‌های مختلف است. عوامل مختلفی مثل وزش باد و طوفان، جریان‌های دریایی، اثر نیروی جاذبه ماه و خورشید روی دریاها، زلزله، آتشفشان و سایر عوامل می‌توانند سبب وقوع امواج یا تغییر در آنها شوند. ارتفاع این امواج می‌تواند از چند سانتی‌متر تا ده‌ها متر باشد. موج پدیده‌ای بسیار غیرخطی و پیچیده بوده که شناسایی و تحلیل آن دشوار است. با وجود اینکه تاکنون

¹ Abnormal waves

دریاهای شمال، ژاپن، خلیج مکزیک، آرام و بالتیک مطالعه شده است.

(Cherneva and Soares, 2014; Liu and Mori, 2000; Veltcheva and Soares, 2012; Chien et al., 2002 and Didenkova, 2011)

موج غیر نرمال یک رویداد گذرا و غیر ایستا است که مطالعه آن با استفاده از روش‌های رایج به دلیل فرض ایستا بودن در فرایند تحلیل در این روش‌ها، دشوار است (Veltcheva and Soares, 2015). در واقع معیارهای ارائه شده برای تشخیص موج غیر نرمال از موج نرمال نیازمند شفاف‌سازی بیشتری است. به عنوان مثال وقتی یک مرتبه بالاتر برای تخمین پروفیل سطح موج فرض می‌شود، این معیارها تغییر می‌کنند، زیرا در مقایسه با تئوری خطی، خاصیت غیرخطی سبب افزایش تیزی موج، افزایش تاج بیشینه، و افزایش ارتفاع موج می‌شود. علاوه بر این تعریف موج غیر نرمال وابسته به زمان اندازه‌گیری رکورد موج است. هرچه مدت زمان اندازه‌گیری بیشتر باشد، احتمال رویت یک موج غیر نرمال افزایش می‌یابد (Dean, 1990). به نظر مؤلفان موج غیرعادی موجی است که از منظر آماری به طور قابل توجهی از الگوی نرمال جامعه آماری تبعیت نکرده و با یک روند متفاوت از سایر داده‌ها تولید شده است.

با بررسی داده‌های تاریخیچه زمانی هواشناسی و ارتفاع امواج یک منطقه خاص می‌توان وقوع امواج غیرعادی را تا حد قابل قبولی پیش‌بینی کرد. هدف از این تحقیق ارائه یک روش ترکیبی برای پیش‌بینی وقوع امواج غیرعادی از روی متغیرهای هواشناسی با استفاده از روش‌های تبدیل موجک⁷ (WT) و ماشین بردار پشتیبان⁸ (SVM) است.

2- روش انجام تحقیق

در شکل 1، روندنمای کلی روش بکار برده شده در این تحقیق برای تشخیص وقوع امواج عادی و غیرعادی ارائه شده است. بدین منظور ابتدا تبدیل موجک روی داده‌های ارتفاع موج مشخصه از نواحی مورد بررسی اعمال شد. با استفاده از این روش داده‌های ارتفاع امواج در دو کلاس غیرعادی (کلاس اول) و کلاس عادی (کلاس دوم) دسته‌بندی می‌شوند.

است تا بتوان از وقوع خسارت‌های هنگفت جانی و مالی جلوگیری کرد. امواج غیرعادی یا غیرنرمال بسیار بزرگ بوده و معمولاً بسیار تیز هستند. این امواج یک پدیده گذرا¹ و غیر ایستا² هستند که در لحظه وقوع می‌توانند بسیار مخرب باشند و سبب آسیب به سازه‌های دریایی و سکوه‌های شناور و کشتی‌ها شوند.

تبدیلات موجک دارای کاربردهای فراوانی در یادگیری ماشین به خصوص پردازش سیگنال است. از تبدیل موجک برای مطالعه زمان-فرکانس توزیع انرژی استفاده می‌شود. این تبدیل سری زمانی داده‌ها را به تعدادی امواج کوچک (موجک³) تجزیه می‌کند. این کار با استفاده از انتخاب یک موجک مادر⁴ انجام می‌شود. تبدیل موجک، توزیع محلی زمان-فرکانس انرژی را با استفاده از ضرب سیگنال در موجک‌ها بدست می‌آورد. در مقایسه با تبدیل فوریه، تبدیل موجک امکان محلی سازی انرژی هم در بازه زمان و هم در بازه مکان را فراهم می‌کند. به همین دلیل برای مطالعه فرایندهای پویا، مثل بررسی امواج، بسیار مناسب است (Massel, 2001; Huang, 2004; Liu and Mori, 2000; Ewans and Buchner, 2008).

Dean (1990) در مطالعه ای مبتنی بر مطالعه یک رکورد موج در 20 دقیقه، بر حسب تئوری موج خطی، موج غیر نرمال را موجی تعریف کرد که ارتفاع آن H بزرگتر از دو برابر ارتفاع موج مشخصه H_s باشد. بعدها نسبت $AI = H_{max}/H_s$ به عنوان شاخص غیرنرمال⁵ موج نام‌گذاری شد. تعریف مشابهی نیز بر حسب تئوری مرتبه دوم ارائه شد. طبق این تعریف در صورتی که شاخص تاج⁶ $CI = H_{crest}/H_s$ از $1/2$ تجاوز کند، موج غیر نرمال قلمداد می‌شود. H_{crest} ارتفاع بزرگترین موج است (Haver, 2005). ترکیبی از $AI > 2$ و $CI > 1.3$ در سال 2000 برای شناسایی موج غیر نرمال نیز پیشنهاد شد (Tomita and Kawamura, 2000).

Clauss (2000) موج غیر نرمال را به صورت $AI > 2$ و $H_{crest} = 0.6H_{max}$ تعریف کرد. موج غیر نرمال در

¹ Transient

² Non-stationary

³ Wavelet

⁴ Mother wavelet

⁵ Abnormality or amplification index

⁶ Crest amplification index

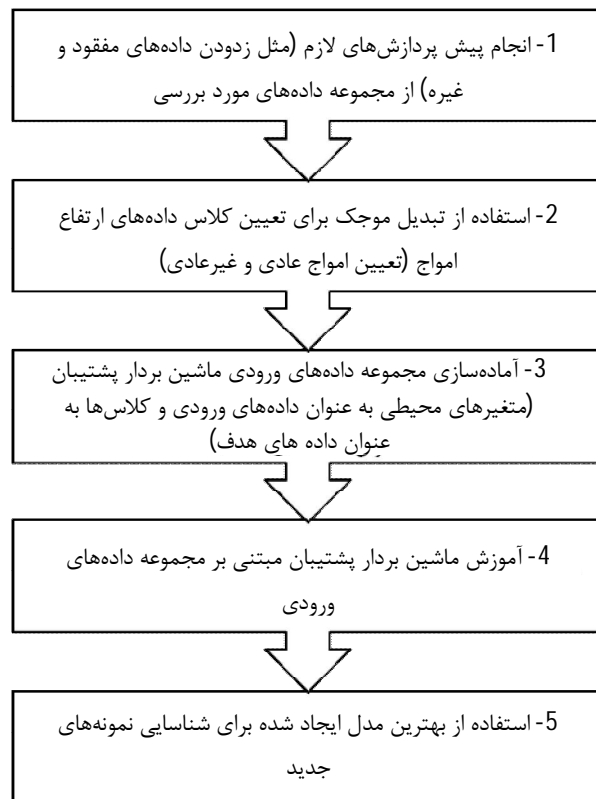
⁷ Wavelet Transform

⁸ Support Vector Machines (SVMs)

مربوط به وقوع یک موج غیرعادی و یا یک موج عادی باشد. در این تحقیق برای مدل‌سازی از داده‌های تاریخی هواشناسی برداشت شده در دو طوفان مشهور استفاده شده است. در بخش 3 جزئیات مجموعه داده‌های مورد بررسی، در بخش‌های 4 و 5 به ترتیب تبدیل موجک و ماشین بردار پشتیبان به طور مختصر معرفی شده است. در بخش 6 نتایج روش‌های بیان شده روی مجموعه داده‌ها، و در بخش 7 نتیجه‌گیری‌های این تحقیق ارائه شده است.

3- منطقه مورد مطالعه و داده‌ها

هدف اصلی تحقیق حاضر تمیز دادن امواج غیرعادی از عادی مبتنی بر متغیرهای هواشناسی با استفاده از الگوریتم‌های محاسبات نرم است. بنابراین به دو دسته داده آموزشی نیازمندیم، یکی مربوط به امواج عادی که بیانگر شرایط نرمال محیطی هستند و دیگری امواج غیرعادی که در شرایط خاص مثل وقوع طوفان رخ می‌دهند. اساس الگوریتم‌های محاسبات نرم استفاده از داده‌های تاریخی جمع‌آوری شده از شرایط سیستم مورد مطالعه است. هرچه داده‌های جمع‌آوری شده دارای وسعت و تنوع بیشتری باشند، در نهایت می‌توان مدلی با انطباق‌پذیری و جامعیت بیشتر ایجاد کرد. داده‌های مورد استفاده در این تحقیق طی سه سال از دو طوفان مشهور (Dean (2007 و Irene (2011 در خلیج مکزیک گردآوری شده است. مسیر گذر این طوفان‌ها در سایت NOAA¹ ترسیم شده است که نشان‌دهنده این است که این مسیر به ایستگاه‌های اندازه‌گیری شماره 41004 و 41041 در خلیج مکزیک بسیار نزدیک است. به همین دلیل در تحقیق حاضر از داده‌های برداشت شده از این دو ایستگاه گردآوری شده است. جزئیات طوفان‌ها و مکان بویه‌ها در جدول 1 و شکل 2 ارائه شده است. از داده‌های تاریخی هواشناسی (جهت باد (WDIR)، سرعت باد (WSPD)، ارتفاع موج مشخصه (WVHT)، فشار سطح دریا (PRES)، دمای هوا (ATMP) و دمای سطح دریا (WTMP)) ثبت شده ساعتی در طول سه سال برای آموزش و آزمون مدل-



شکل 1 روندنمای روش بکار برده شده در این تحقیق برای شناسایی امواج عادی و غیرعادی

سپس از خروجی‌های تبدیل موجک به عنوان داده‌های هدف در ماشین بردار پشتیبان استفاده شده است. ماشین بردار پشتیبان یکی از روش‌های یادگیری با نظارت است که از آن برای مسائل دسته‌بندی و رگرسیون استفاده می‌کنند. این روش دارای کاربردهای فراوانی در علوم نرم است. مدل‌سازی با استفاده از ماشین بردار پشتیبان، نیازمند آموزش است. آموزش ماشین بردار پشتیبان نیازمند تعیین پارامترها و مجموعه داده‌های آموزشی است که خود به داده‌های ورودی و هدف تقسیم‌بندی می‌شود. در تمامی مدل‌های ایجاد شده داده‌های هواشناسی (جهت باد، سرعت باد، ارتفاع موج مشخصه، فشار سطح دریا، دمای هوا و دمای سطح دریا) به عنوان ورودی و کلاس‌های امواج (غیرعادی و عادی) به عنوان خروجی در نظر گرفته شده‌اند. پس از تهیه مدل‌ها می‌توان از آن‌ها برای پیش‌بینی وقوع امواج غیرعادی استفاده کرد، بدین صورت که داده‌های هواشناسی جدید به عنوان ورودی به مدل داده شده، مدل تعیین می‌کند که این داده‌ها می‌تواند

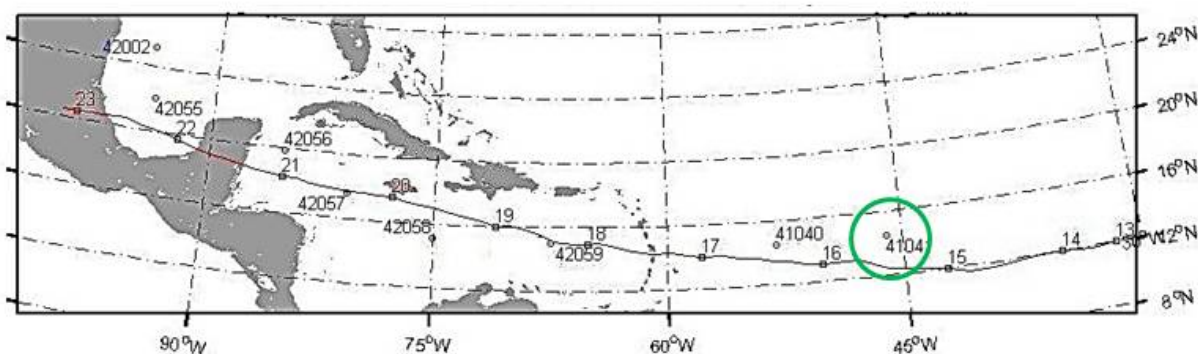
¹ <http://www.ndbc.noaa.gov>

داده‌های آموزشی و داده‌های آزمون ارائه شده است.

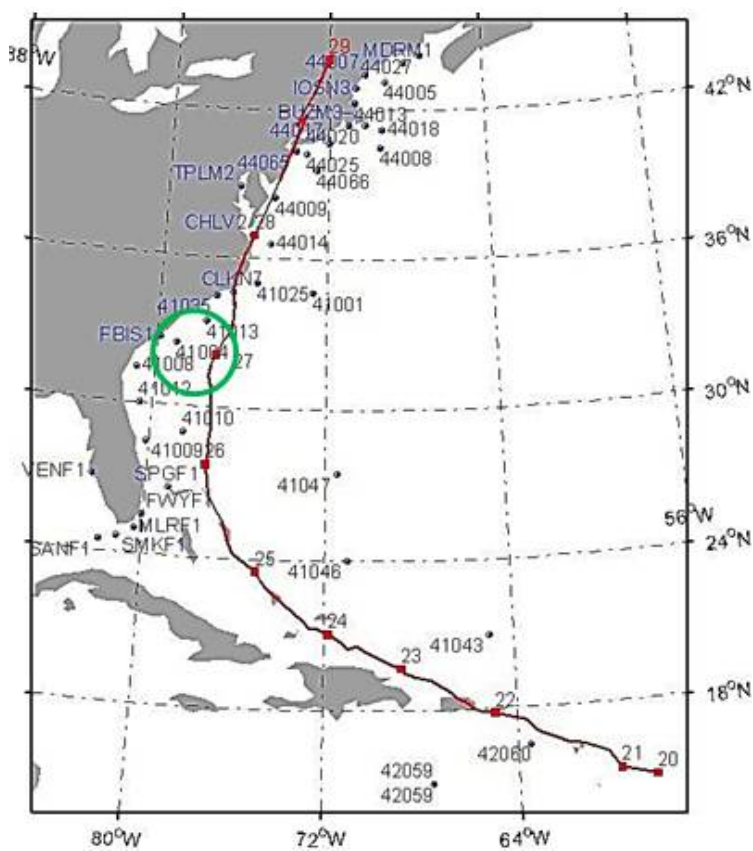
ها استفاده شده است. در جدول 2 مشخصه‌های آماری متغیرهای هواشناسی و ارتفاع امواج مورد بررسی برای

جدول 1 مشخصات بویه‌ها و طوفان‌های مورد بررسی (<http://www.ndbc.noaa.gov>)

ردیف	طوفان	بازه وقوع	ایستگاه	مکان ایستگاه	بازه اندازه‌گیری
1	Dean	13 تا 23 آگوست 2007	41041	14/329 N 46/082 W (14°19'43" N 46°4'55" W)	2005 تا 2007
2	Irene	20 تا 28 آگوست 2011	41004	32/501 N 79/099 W (32°30'2" N 79°5'58" W)	2009 تا 2011



الف) مسیر عبور طوفان Dean (2007)



ب) مسیر عبور طوفان Irene (2011)

شکل 2 مکان بویه‌ها (مشخص شده با علامت دایره) و مسیر عبور طوفان‌های مورد بررسی (<http://www.ndbc.noaa.gov>)

جدول 2 مشخصه‌های آماری متغیرهای هواشناسی مورد بررسی

الف) طوفان (Dean (2007)

مجموعه داده آزمایشی (3809 نمونه)				مجموعه داده آموزشی (4835 نمونه)				متغیر
انحراف معیار	میانگین	بیشترین	کمترین	انحراف معیار	میانگین	بیشترین	کمترین	
33/34	86/47	355	0/0	33/05	86/64	359	0/0	جهت باد (درجه)
1/99	6/97	18/9	0/0	1/95	7/02	19/3	0/1	سرعت باد (متر بر ثانیه)
0/49	1/99	6/87	0/96	0/49	2/00	6/7	0/92	ارتفاع موج مشخصه (متر)
1/94	1014/27	1020/8	1005/8	1/96	1014/33	1020/5	1005/7	فشار سطح دریا (هکتو پاسکال)
1/00	26/43	29/5	21/7	1/02	26/40	29/3	22/2	دمای هوا (سلسیوس)
0/89	27/13	30	25	0/92	27/10	29/8	25	دمای سطح دریا (سلسیوس)

ب) طوفان (Irene (2011)

مجموعه داده آزمایشی (3 نمونه)				مجموعه داده آموزشی (3 نمونه)				متغیر
انحراف معیار	میانگین	بیشترین	کمترین	انحراف معیار	میانگین	بیشترین	کمترین	
97/87	181/18	360	1	97/77	180/71	360	1	جهت باد (درجه)
3/12	6/45	21/8	0	3/12	6/45	21/1	0	سرعت باد (متر بر ثانیه)
0/65	1/25	6/29	0/28	0/66	1/25	7/48	0/29	ارتفاع موج مشخصه (متر)
5/82	1017/19	1034/7	992/7	5/82	1017/19	1035/1	994/9	فشار سطح دریا (هکتو پاسکال)
6/55	20/70	30/7	0/5	6/55	20/70	30/9	0/5	دمای هوا (سلسیوس)
4/39	23/03	32/2	14/3	4/39	23/03	31/5	14/3	دمای سطح دریا (سلسیوس)

(Veltcheva and Soares, 2015):

$$x(t) = \int_{-\infty}^{\infty} c_n \hat{\psi}_n dt \quad (2)$$

که $\{\hat{\psi}_n\}$ توابع پایه متعامد برای $\{\psi_n\}$ است، به گونه‌ای که $\langle \psi_n, \hat{\psi}_m \rangle = \delta_{mn}$. تابع دلتای کرونگر² است. معادله (2)، معکوس تبدیل یا تجزیه نامیده می‌شود. توابع $\{\psi_n(t)\}$ مبنای تجزیه را شکل می‌دهند. روش‌های مختلفی برای تجزیه یک سیگنال وجود دارد.

بر خلاف تبدیل فوریه، تبدیل موجک از یک روش متفاوت برای تجزیه سیگنال به اجزای تشکیل دهنده آن استفاده می‌کند. تبدیل موجک (WT) سیگنال $x(t)$ بوسیله ضرب داخلی تعریف می‌شود (Veltcheva and Soares, 2015):

$$WT(\tau, b) = \langle x(t), g_{\tau b}(t) \rangle = \int_{-\infty}^{\infty} x(t) g_{\tau b}^*(t; \tau, b) dt \quad (3)$$

در اینجا کاراکتر * بیانگر مزدوج مختلط بوده و $g_{\tau b}(t)$ از موجک مادر $g(t)$ حاصل می‌شود (Veltcheva and Soares, 2015):

$$g_{\tau b}(t; \tau, b) = \frac{1}{\sqrt{b}} g\left(\frac{t-\tau}{b}\right) \quad (4)$$

4- تبدیل موجک

تبدیل موجک به عنوان روشی جایگزین برای تبدیل فوریه معرفی شد و هدف از ارائه آن رفع مشکلات تبدیل فوریه بود (Veltcheva and Soares, 2015). استخراج اطلاعات در خصوص یک پدیده با استفاده از داده‌های جمع‌آوری شده از رفتار آن و مقایسه این داده‌ها با مجموعه‌ای از توابع شناخته شده، هدف اصلی پردازش سیگنال است. این مقایسه برای یک سیگنال پیوسته با استفاده از ضرب داخلی سیگنال $x(t)$ و توابع $\psi_n(t)$ انجام می‌شود (Veltcheva and Soares, 2015):

$$c_n = \langle x, \psi_n \rangle = \int_{-\infty}^{\infty} x(t) \psi_n^*(t) dt \quad (1)$$

در اینجا کاراکتر * بیانگر مزدوج مختلط است. ضرب داخلی بیانگر میزان شباهت سیگنال $x(t)$ و تابع $\psi_n(t)$ است. هرچه میزان شباهت سیگنال $x(t)$ به تابع $\psi_n(t)$ بیشتر باشد، حاصل ضرب داخلی بیشتر خواهد بود. رابطه (1) تبدیل¹ نامیده می‌شود. در صورتی که تابع $\{\psi_n(t)\}$ متعامد باشد، سیگنال $x(t)$ بوسیله رابطه (2) بازیابی می‌شود

² Kronecker delta function

¹ Transformation

2001; Chien et al., 2002; Veltcheva and Soares, 2015). این موجک شامل یک جمله گوسی و یک جمله مدولاسیون با فرکانس c بوده و به صورت رابطه (7) است (Veltcheva and Soares, 2015):

$$g(t) = \exp\left(-\frac{1}{2}t^2\right) \exp(ict) \quad (7)$$

بر این اساس رابطه تبدیل موجک پیوسته با موجک مادر مورلت به صورت رابطه (8) تعریف می‌شود (Veltcheva and Soares, 2015):

$$g_{\tau b}(t) = \frac{1}{\sqrt{b}} \exp\left(-\frac{1}{2}\left(\frac{t-\tau}{b}\right)^2\right) \exp\left(ic\frac{t-\tau}{b}\right) \quad (8)$$

انرژی کل در تبدیل موجک به صورت رابطه (9) است (Veltcheva and Soares, 2015):

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = C^{-1} \int_0^{\infty} \int_0^{\infty} |WT(\tau, b)|^2 d\tau db \quad (9)$$

مربع تبدیل موجک، scalogram یا طیف انرژی موجک نامیده می‌شود (Veltcheva and Soares, 2015):

$$E(\tau, b) = |WT(\tau, b)|^2 \quad (10)$$

طیف‌های مختلف انرژی موجک و چگالی‌های طیفی با استفاده از رابطه (10) تعریف می‌شوند. چگالی انرژی زمان-مقیاس به صورت رابطه (11) تعریف می‌شود (Veltcheva and Soares, 2015):

$$E_1(\tau, b) = \frac{|WT(\tau, b)|^2}{b} \quad (11)$$

چگالی انرژی محلی بوسیله انتگرال‌گیری از رابطه (11) نسبت به پارامتر مقیاس b به صورت رابطه (12) حاصل می‌شود:

$$E_2(\tau) = C^{-1} \int_0^{\infty} E_1(\tau, b) \frac{db}{b} \quad (12)$$

$E_2(\tau)$ طیف انرژی بر واحد فرکانس است. انتگرال‌گیری از معادله (11) نسبت به زمان τ طیف موجک عمومی $E_3(b)$ را نتیجه می‌دهد (Veltcheva and Soares, 2015):

$$E_3(b) = \int_0^{\infty} E_1(\tau, b) d\tau \quad (13)$$

انرژی کلی سیگنال $x(t)$ به صورت رابطه (14) تعریف می‌شود (Veltcheva and Soares, 2015):

$$E(\tau) = C^{-1} \int_0^{\infty} E_3(b) \frac{db}{b} \quad (14)$$

که در آن τ و b به ترتیب پارامترهای انتقال¹ و مقیاس² هستند. τ اطلاعات زمانی تبدیل را در برداشته و بیانگر مکان موجک است، هنگامی که در طول سیگنال جابجا می‌شود. در تبدیل موجک به طور مستقیم پارامتر فرکانس وجود ندارد، اما پارامتر مقیاس وجود دارد که به طور معکوس با فرکانس ارتباط دارد. در واقع $b = 1/f$. پارامتر مقیاس به عنوان یک اپراتور ریاضی، سیگنال را منقبض ($b < 1$) و یا منبسط ($b > 1$) می‌کند. در واقع مقیاس‌های بزرگ متناظر با فرکانس‌های پایین، و مقیاس‌های کوچک متناظر با فرکانس‌های بالا هستند.

ضریب موجک $WT(\tau, b)$ بیانگر میزان همبستگی بین موجک و بخش موضعی سیگنال است. مقادیر بزرگ $WT(\tau, b)$ بیانگر این مطلب است که سیگنال مورد بررسی دارای یک جزء اصلی فرکانس متعلق به مقیاس داده شده است. در این صورت موجک مقیاس شده شبیه سیگنال مورد مطالعه خواهد بود. با تغییر مقادیر پارامترهای τ و b ، می‌توان تغییرات دامنه بر حسب زمان و مقیاس را ایجاد کرد.

موجک مادر باید دارای میانگین صفر باشد. در نتیجه سیگنال اصلی را می‌توان با استفاده از ضرایب تبدیل موجک $WT(\tau, b)$ بوسیله معکوس تبدیل موجک به صورت رابطه (5) محاسبه کرد (Veltcheva and Soares, 2015):

$$x(t) = \frac{1}{C} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} WT(\tau, b) b^{-2} g_{\tau b}(t; \tau, b) d\tau db \quad (5)$$

که:

$$C = \int_{-\infty}^{\infty} (\omega^{-1} |G(\omega)|^2) d\omega < \infty \quad (6)$$

$G(\omega)$ تبدیل فوریه تابع $g(t)$ است. این شرط به عنوان شرط پذیرش³ شناخته می‌شود. برگشت‌پذیر بودن تبدیل و توانایی بازسازی سیگنال $x(t)$ از روی تبدیل موجک پیوسته آن به این شرط وابسته است.

پس از اینکه موجک پایه انتخاب شد، می‌توان از آن برای تحلیل داده‌ها استفاده کرد. برای تحلیل داده‌های موج، استفاده از موجک مورلت⁴ با توجه به ماهیت نوسانی آن می‌تواند مفید باشد (Liu and Mori, 2000; Massel,

¹ Translation parameter

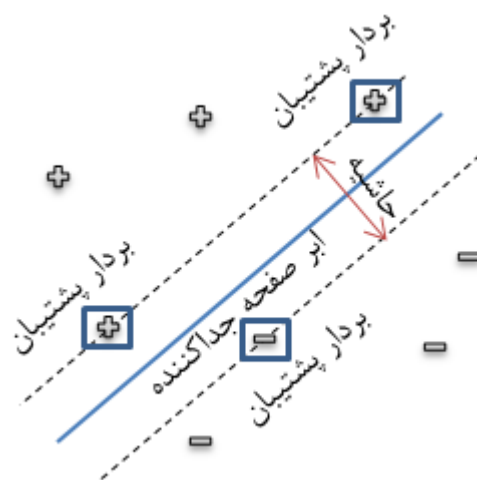
² Scale dilation parameter

³ Admissibility condition

⁴ Morlet

5- ماشین بردار پشتیبان دسته‌بندی کننده

ماشین‌های بردار پشتیبان (SVM) یکی از روش‌های دسته‌بندی کننده دو کلاسه با یادگیری با ناظر است که توسط Vapnik, (2000) توسعه یافت. می‌توان از SVM برای مسائل دسته‌بندی (که خروجی یک دسته است) یا مسائل رگرسیون (که خروجی یک مقدار عددی است) استفاده کرد. این روش، داده‌ها را با یافتن بهترین ابرصفحه‌هایی¹ که تمامی داده‌های یک کلاس را از داده‌های کلاس دیگر مجزا می‌کنند، دسته‌بندی می‌کند. در واقع ایده اصلی ماشین‌های بردار پشتیبان ترسیم ابرصفحه‌هایی در فضا است که عمل تمایز نمونه‌های مختلف داده را به طور بهینه انجام دهند. بهترین ابر صفحه، به معنای صفحه‌ای است که دارای بیشترین حاشیه از دو کلاس است. حاشیه به معنای ورقه‌ای است که بیشترین عرض موازات ابرصفحه جداکننده را دارا است، به نحوی که درون آن هیچ داده‌ای واقع نشده باشد. بردارهای پشتیبان به نقاطی گفته می‌شود که دارای کمترین فاصله با ابرصفحه جداکننده هستند. این نقاط روی مرز ورقه واقع شده‌اند. شکل 3 بیانگر این تعاریف است. داده‌های نشان داده شده با کاراکتر '+' و '-' به ترتیب مشخص کننده دسته 1 و دسته -1 هستند. در ادامه به طور مختصر روش SVM معرفی شده است. جزئیات این روش را می‌توان در (Scholkopf and Smola (2002) و Scholkopf (1999) یافت.



شکل 3 مفاهیم ماشین‌های بردار پشتیبان

بردارهای x_i برای $1 \leq i \leq N_x$ به عنوان ورودی، و بردارهای برچسب دودویی $y_i \in \{-1, 1\}$ به عنوان خروجی متناظر با آن در نظر گرفته می‌شوند (بنابراین مقدار y_i وقتی که x_i در کلاس 1 واقع باشد، برابر 1، و وقتی که x_i در کلاس 2 واقع باشد برابر -1 است). N_x تعداد نمونه‌ها است. $\phi(x_i)$ به عنوان بردارهای تشابه در فضای ویژگی، و $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ به عنوان تابع کرنل که دلالت بر ضرب داخلی در فضای ویژگی دارد، تعریف می‌شوند.

مسئله بهینه‌سازی در SVM برای یک مسئله حاشیه هموار نسبت به $y_i(w \cdot x + b) = 1 - \xi_i$ و $\xi_i \geq 0$ ، به صورت رابطه (15) تعریف می‌شود (Vapnik, 2000):

$$\min_{w,b} = \left\{ \frac{1}{2} \|w\|^2 + C \sum_i \xi_i \right\} \quad (15)$$

که بردار نرمال ابرصفحه جداکننده در فضای ویژگی، و $C > 0$ یک پارامتر تنظیم کننده است که وظیفه کنترل کردن دسته‌بندی ناصحیح را بر عهده دارد. با استفاده از لاگرانژین رابطه (15) فرم دوگان حاصل می‌شود (Vapnik, 2000):

$$w(\alpha) = \max_{\alpha} \left\{ \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right\} \quad (16)$$

که در آن $0 \leq \alpha \leq C$ است. این یک مسئله بهینه‌سازی درجه دوم است که می‌توان آن را به خوبی با الگوریتم‌هایی مثل بهینه‌سازی کمینه ترتیبی² (Scholkopf et al., 1998) حل کرد. در واقع حین بهینه‌سازی بسیاری از α_i ها به سمت صفر میل می‌کنند و x_i های باقی‌مانده متعلق به $\alpha_i > 0$ بردارهای پشتیبان نامیده می‌شوند. در اینجا برای سادگی فرض می‌کنیم تمامی داده‌های غیر بردار پشتیبان حذف شده‌اند. بنابراین اکنون N_x بیانگر تعداد بردارهای پشتیبان و برای تمامی‌ها $\alpha_i > 0$ است. در نتیجه، بردار نرمال صفحه جداکننده w به صورت رابطه (17) محاسبه می‌شود (Vapnik, 2000):

$$w = \sum_{i=1}^{N_x} \alpha_i y_i x_i \quad (17)$$

چون $\phi(x_i)$ به صورت ضمنی تعریف شد، w فقط در فضای ویژگی وجود دارد و نمی‌تواند به صورت مستقیم محاسبه شود. در عوض، دسته‌بندی $f(q)$ برای یک بردار نمونه جدید q فقط می‌تواند بوسیله محاسبه تابع کرنل q با هر

² Sequential Minimal Optimization

¹ Hyperplane

شعاع است. در صورتی که مقدار γ بزرگ انتخاب شود، مدل دچار بیش برآزش می‌شود. در روش ماشین‌های بردار پشتیبان نیز همانند هر مدل یادگیری با ناظر، ابتدا باید مدل را با استفاده از داده‌های آموزشی، آموزش داد. پس از آموزش می‌توان از آن برای دسته‌بندی نمونه‌های جدید استفاده کرد. برای دستیابی به بهترین مدل (بالاترین دقت)، می‌توان از کرنل‌های مختلف استفاده کرد و با انتخاب مناسب پارامترهای آن‌ها دقت را افزایش داد.

5-1- معیارهای ارزیابی عملکرد

در این تحقیق برای بررسی عملکرد مدل‌های ماشین بردار پشتیبان ایجاد شده از معیارهای زیر استفاده شده است:
- جزر میانگین مربعات خطا ($RMSE$):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - x'_i)^2} \quad (20)$$

در این رابطه، x_i داده‌های اندازه‌گیری شده (مقادیر واقعی)، x'_i داده‌های پیش‌بینی شده (خروجی‌های شبکه)، و n تعداد داده‌ها است.

- ماتریس در هم ریختگی²: این ماتریس یکی از روش‌های رایج بررسی صحت عملکرد الگوریتم‌هایی است که از یادگیری نظارت شده استفاده می‌کنند. ماتریس در هم ریختگی یک ماتریس مربعی $N \times N$ است که N بیانگر تعداد کلاس‌ها است. در این ماتریس روی محور افقی کلاس‌های واقعی، و روی محور عمودی کلاس‌های تخمینی الگوریتم نشان داده می‌شود. اعداد روی قطر اصلی ماتریس نشان دهنده تعداد کلاس بندی‌های انجام شده به صورت صحیح است. بنابراین در صورتی که تمام اعداد قرار گرفته روی درایه‌های غیر قطر اصلی صفر باشند، الگوریتم دارای دقت حداکثر است.

- دقت: پارامتر دقت به صورت نسبت تعداد نمونه‌های کلاس‌بندی شده به صورت صحیح بر تعداد کل نمونه‌های مجموعه داده تعریف می‌شود.

- منحنی ROC³: این منحنی نیز یکی از شاخص‌های

بردار پشتیبان صورت گیرد (Vapnik, 2000):

$$f(q) = \text{sign} \left(\sum_{i=1}^{N_x} \alpha_i y_i \cdot K(q, x_i) + d \right) \quad (18)$$

در این رابطه پارامتر d انحراف ابر صفحه در طول بردار نرمال خود است که در طی فرایند آموزش تعیین می‌شود.

یکی از مراحل مهم در تعیین عملکرد ماشین‌های بردار پشتیبان برای دستیابی به یک دسته‌بندی هرچه صحیح‌تر، انتخاب مناسب تابع کرنل است. این تابع با توجه به ماهیت داده‌های مورد مطالعه انتخاب می‌شود. بنابراین درک درست از داده‌ها می‌تواند به انتخاب صحیح تابع کرنل کمک کند. زمانی که کاربر شناخت کافی از داده‌ها نداشته و قادر به انتخاب مناسب تابع کرنل نباشد، استفاده از روش‌های خودکار مثل اعتبارسنجی متقابل¹ و یا استفاده ترکیبی از دسته‌بندی کننده‌ها با کرنل‌های مختلف می‌تواند راهگشا باشد. SVM برای تفکیک داده‌های غیرخطی، از کرنل‌های مختلف استفاده می‌کند. با استفاده از این کرنل‌ها، داده‌ها به فضایی با ابعاد بیشتر نگاشت می‌شوند تا بتوان آن‌ها را در این فضای جدید به صورت خطی تفکیک کرد. تعداد زیادی تابع کرنل در SVM وجود دارد. توابع پایه‌ای شعاعی گوسی و چندجمله‌ای‌ها دو تابع معروف و پرکاربرد در SVM به شمار می‌آیند (Scholkopf and Smola, 2002; Scholkopf, 1999). کرنل‌ها به دو دسته خطی و غیرخطی تقسیم می‌شوند. کرنل‌های خطی داده‌ها را در فضای اصلی به صورت خطی، و کرنل‌های غیرخطی داده‌ها را با استفاده از مرزهای غیرخطی جدا می‌کنند. آموزش مدل با استفاده از کرنل‌های خطی سریعتر از کرنل‌های غیرخطی انجام می‌شود. همچنین استفاده از کرنل‌های خطی احتمال بیش برآزش مدل‌های ایجاد شده را کاهش می‌دهد. کرنل گوسی یکی از محبوب‌ترین کرنل‌های مورد استفاده در SVM است که بینهایت مشتق‌پذیر بوده، در نتیجه می‌توان داده‌ها را به یک فضای ویژگی با ابعاد بینهایت نگاشت کرد. در این تحقیق از کرنل تابع شعاعی گوسی استفاده شده است (Taylor and Cristianini, 2004):

$$RBF(x, z) = \exp(-\gamma \|x - z\|^2) \quad (19)$$

در این رابطه پارامتر γ مقداری مثبت بوده که کنترل کننده

2. Confusion matrix

3. Receiver Operating Characteristic (ROC)

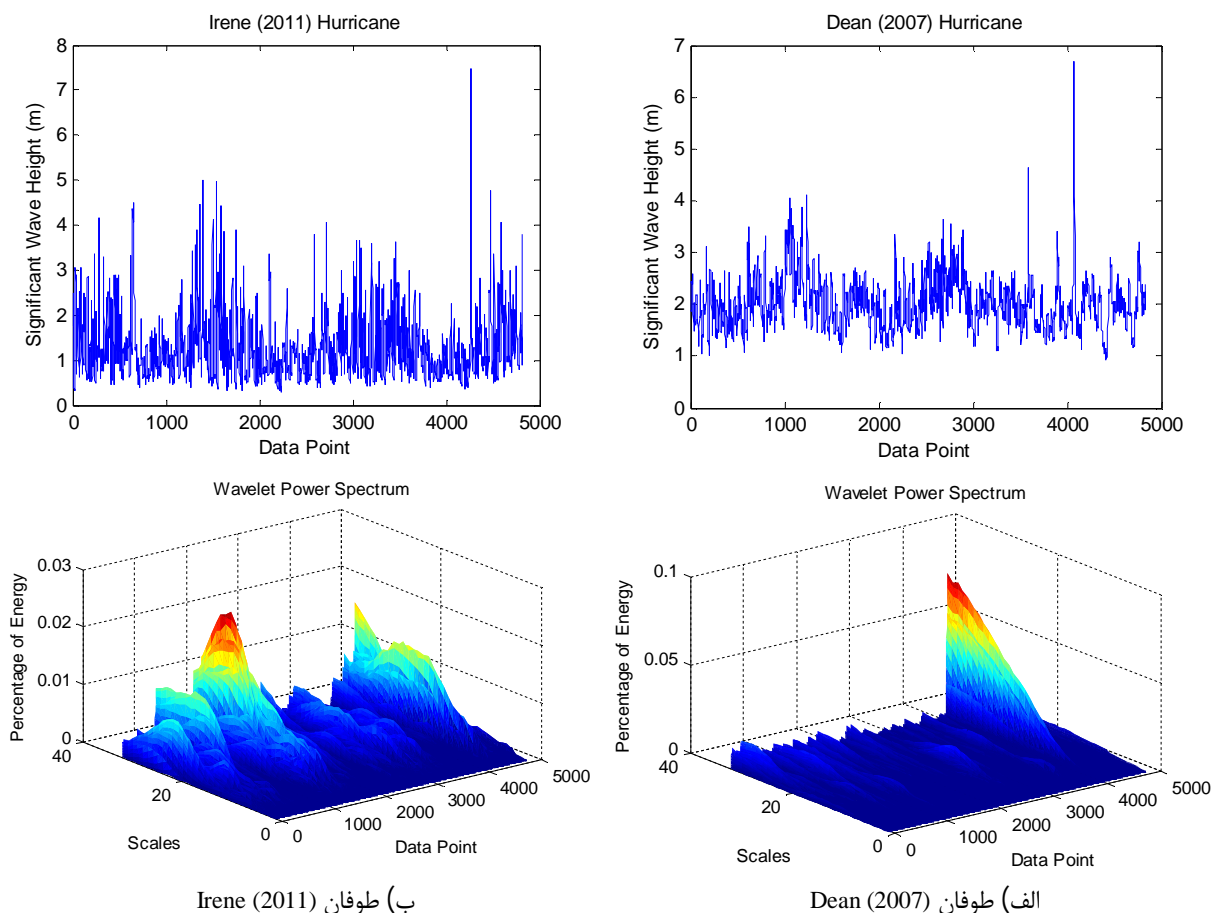
1. Cross validation

برای پیش‌بینی تولید امواج غیرعادی ابتدا باید امواج عادی از امواج غیرعادی تمیز داده شوند. این کار با استفاده از تبدیل موجک داده‌های ارتفاع موج مشخصه قابل انجام است. در شکل‌های 4 و 5 طیف انرژی موجک (رابطه 10) مبتنی بر تبدیل موجک پیوسته امواج مورد مطالعه به ترتیب برای مجموعه داده‌های آموزشی و آزمایشی ارائه شده است. موجک مورلت تعریف شده بوسیله رابطه (7) به عنوان موجک مادر انتخاب شده است. با توجه به نتایج تحقیق (Veltcheva and Soares (2015)، امواج غیرعادی در محدوده طیف با انرژی زیاد ظاهر می‌شوند. بنابراین در شکل‌های 4 و 5 در نواحی از طیف که انرژی به طور قابل ملاحظه‌ای نسبت به سایر نقاط زیادتر باشد، به معنای وقوع امواج غیرعادی در آن نواحی است. این نواحی به صورت پیک در نمودار طیف ظاهر می‌شوند.

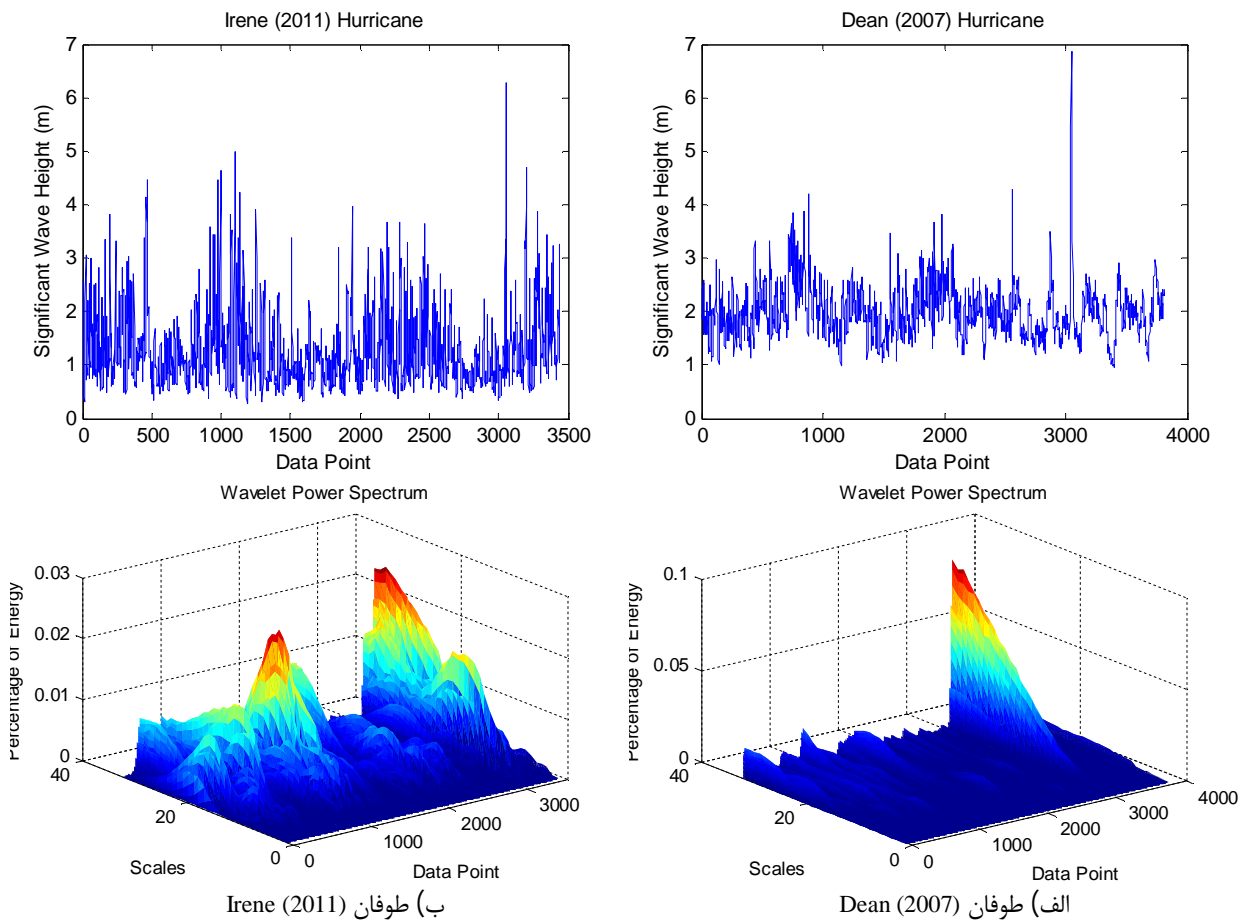
ارزیابی عملکرد الگوریتم‌هایی است که از یادگیری نظارت شده و یادگیری بدون ناظر استفاده می‌کنند. در این نمودار هر چه نقاط به سمت بالا و چپ نزدیک‌تر باشند، مدل ایجاد شده دارای دقت بیشتری بوده و به حالت ایده‌آل خود نزدیک‌تر است. مختصات نقطه‌ای (0 و 1) حالت ایده‌آل است. این نقطه بدین معنا است که خروجی‌های مدل با مقادیر واقعی تطابق کامل دارد. هرچه مساحت زیر نمودار منحنی‌های ROC بیشتر باشد، مدل ایجاد شده دارای عملکرد بهتری است.

6- نتایج

در این بخش عملکرد روش ترکیبی WT-SVM برای پیش‌بینی تولید امواج عادی و غیرعادی با استفاده از متغیرهای هواشناسی بررسی شده است. برای انجام تمامی محاسبات و تحلیل‌ها از نرم‌افزار متلب استفاده شده است.



شکل 4 امواج مورد مطالعه و طیف انرژی موجک پیوسته آن‌ها برای مجموعه داده‌های آموزشی



شکل 5 امواج مورد مطالعه و طیف انرژی موجک پیوسته آن‌ها برای مجموعه داده‌های آزمایشی

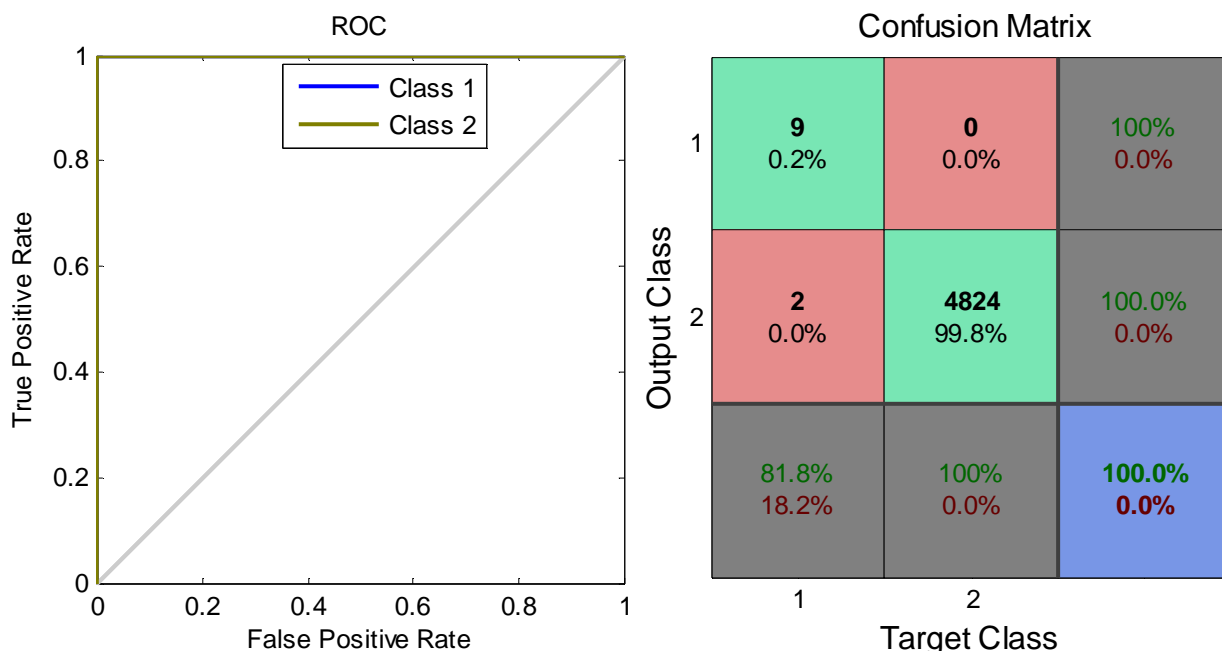
خطا به 0 نزدیک‌تر باشد، مدل طراحی شده از کارایی مطلوب‌تری برخوردار است. همانطور که از نتایج این جدول مشخص است، کارایی مدل‌های ایجاد شده روی مجموعه داده‌های آموزشی و آزمایشی مناسب بوده است. علت حصول دقت بالای مدل‌ها، مناسب بودن فاز آموزش هر مدل بوده است. همچنین به علت اختلاف زیاد داده‌های موج غیرعادی از داده‌های نرمال، مدل‌ها توانسته‌اند با دقت مطلوبی شناسایی موج غیرعادی را انجام دهند. در شکل‌های 6 و 7، نمودارهای عملکرد مدل بهینه طراحی شده برای مجموع داده‌های مورد بررسی ارائه شده است. نمودارها شامل موارد زیر است:

- نمودار در هم ریختگی داده‌های آموزشی و داده‌های آزمایشی. در این شکل تعداد نمونه‌هایی که به صورت صحیح دسته‌بندی شده‌اند، روی قطر اصلی ماتریس و در خانه‌های سبز رنگ قرار می‌گیرند.

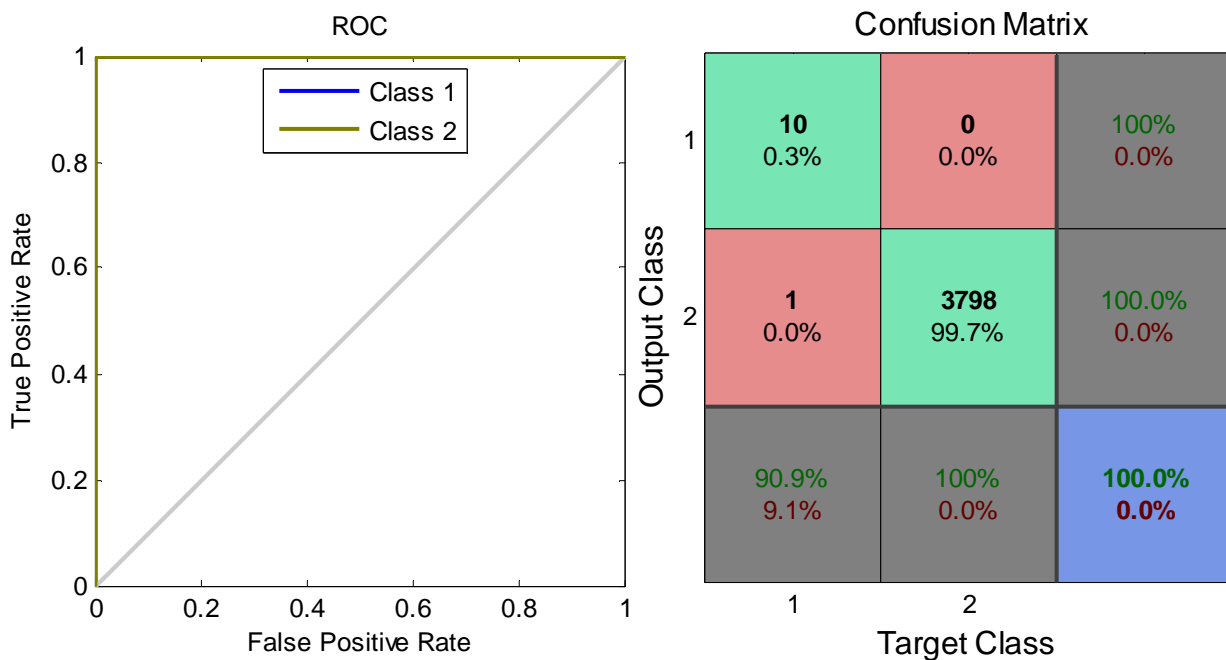
در این تحقیق، امواج غیرعادی با کلاس 1، و امواج عادی با کلاس 2 معرفی شده‌اند. پس از مشخص شدن کلاس داده‌ها، می‌توان ماشین بردار پشتیبان را طراحی کرد. برای ایجاد مدل‌ها، باید آن‌ها را تحت پارامترهای موثر بر عملکرد آن‌ها تحت آموزش و یادگیری قرار داد و در نهایت بهینه‌ترین پارامترها (که منجر به بهترین پاسخ می‌شوند) را انتخاب کرد. برای مدل‌سازی داده‌های امواج طوفان‌های Irene و Dean دو مدل مختلف ماشین بردار پشتیبان طراحی شده است. در هر دو مدل از تابع کرنل گوسی با $\gamma = 20$ و $\sigma^2 = 0.44$ استفاده شده است. هر مدل دارای 5 متغیر ورودی (جهت باد (WDIR)، سرعت باد (WSPD)، فشار سطح دریا (PRES)، دمای هوا (ATMP) و دمای سطح دریا (WTMP)) و یک متغیر خروجی (کلاس امواج) است. در جدول 3 نتایج عملکرد مدل‌های ایجاد شده روی مجموعه داده‌های آموزشی و آزمایشی ارائه شده است. هرچه مقدار پارامتر دقت به به 100 (بر حسب درصد)، و

جدول 3 نتایج مدل‌سازی پیش‌بینی امواج غیرعادی طوفان (Dean (2007) و Irene (2011) با استفاده از ماشین بردار پشتیبان

مجموعه داده آزمایشی		مجموعه داده آموزشی		طوفان
خطا	دقت (درصد)	خطا	دقت (درصد)	
0/16	100	0/16	100	Dean (2007)
0/18	99/90	0/17	99/90	Irene (2011)

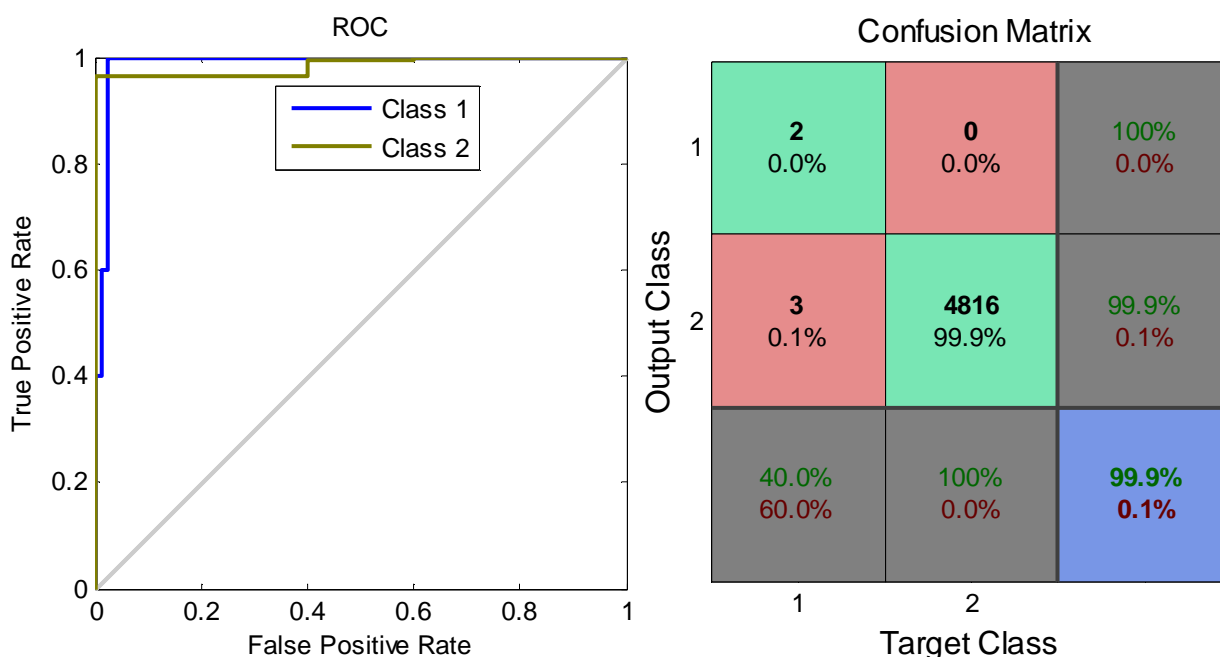


الف) نمودار در هم ریختگی و منحنی ROC داده‌های آموزشی

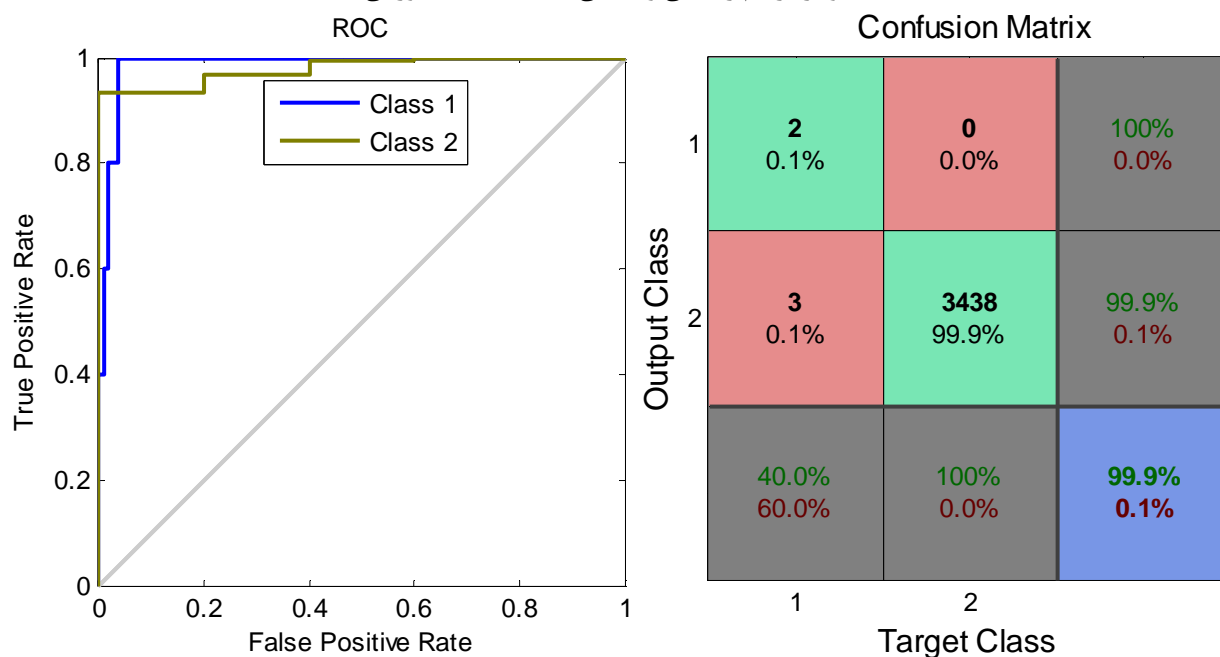


ب) نمودار در هم ریختگی و منحنی ROC داده‌های آزمایشی

شکل 6 نمودارهای عملکرد مدل بهینه طراحی شده برای مجموعه داده طوفان (Dean (2007)



الف) نمودار در هم ریختگی و منحنی ROC داده‌های آموزشی



ب) نمودار در هم ریختگی و منحنی ROC داده‌های آزمایشی

شکل 7 نمودارهای عملکرد مدل بهینه طراحی شده برای مجموعه داده طوفان (Irene (2011)

این شکل، خطوط رنگی در هر محور بیانگر منحنی‌های ROC برای هر کلاس است. هرچه این خطوط به سمت چپ گوشه بالا نمودار متمرکز شده باشند، بدین معنا است که عملکرد مدل ایجاد شده مناسب‌تر است. در جدول 4 خلاصه عملکرد بهترین مدل‌های ایجاد شده در

خانه واقع شده در پایین و سمت راست ماتریس (مشخص شده با رنگ آبی) بیانگر میزان دقت کلی مدل است. به طور کلی سطر سوم و ستون سوم هر شکل، بیانگر میزان دقت و خطای کلاس‌بندی بر حسب درصد است. - منحنی ROC داده‌های آموزشی و داده‌های آزمایشی. در

پشتیبان استفاده شد. بدین منظور می‌توان از سایر الگوریتم‌های کلاس‌بندی مثل شبکه‌های عصبی مصنوعی، روش k-تا نزدیک‌ترین همسایه و درختان تصمیم نیز استفاده کرد. در جدول 5 نتایج بدست آمده با استفاده از کلاس‌بندی به روش شبکه عصبی مصنوعی تغذیه رو به جلو دو لایه (Lavine and Blank, 2009) به ازای تعداد 10 نرون ارائه شده است. با توجه به نتایج این جدول می‌توان گفت روش شبکه عصبی نیز توانسته است با دقت مطلوبی امواج غیرعادی را از امواج عادی تمیز دهد. بنابراین می‌توان از این روش نیز بجای ماشین بردار پشتیبان برای کلاس‌بندی داده‌های موج استفاده کرد.

تشخیص امواج غیرعادی (سهمگین) از عادی (نرمال) بر اساس نتایج حاصل از ماتریس در هم ریختگی نشان داده شده است. با توجه به نتایج این جدول می‌توان نتیجه گرفت که به طور کلی دقت مدل‌ها در تشخیص امواج عادی از امواج غیرعادی بیشتر است. علت این است که تعداد نمونه‌های آموزشی امواج عادی بیشتر از امواج غیرعادی بوده است و مدل توانسته همبستگی بیشتری بین امواج عادی پیدا کند. پس از ایجاد مدل‌ها، می‌توان از آن‌ها به عنوان یک ابزار مناسب برای پیش‌بینی وقوع امواج غیرعادی با استفاده از متغیرهای هواشناسی استفاده کرد. در این تحقیق برای کلاس‌بندی امواج از روش ماشین بردار

جدول 4 خلاصه عملکرد بهترین مدل‌های ایجاد شده برای امواج مورد بررسی در تشخیص امواج غیرعادی از عادی (نرمال)

تعداد امواج غیر نرمال	تعداد امواج غیر نرمال	تعداد امواج غیر نرمال	تعداد امواج غیر نرمال	تعداد امواج غیر نرمال	تعداد امواج غیر نرمال	مجموعه داده طوفان
شناسایی شده به	شناسایی شده به	شناسایی شده به	شناسایی شده به	شناسایی شده به	شناسایی شده به	
عنوان غیر نرمال	عنوان غیر نرمال	عنوان غیر نرمال	عنوان غیر نرمال	عنوان غیر نرمال	عنوان غیر نرمال	
2	9	2	0	4828	4837	آموزشی Dean
1	10	1	0	3798	3808	آزمایشی
3	2	3	0	4816	4818	آموزشی Irene
3	2	3	0	3438	3440	آزمایشی

جدول 5 نتایج مدل‌سازی پیش‌بینی امواج غیرعادی طوفان (Dean (2007) و Irene (2011) با استفاده از شبکه عصبی دسته‌بندی کننده

مجموعه داده آموزشی		مجموعه داده آزمایشی		طوفان
خطا	دقت (درصد)	خطا	دقت (درصد)	
0/14	100	0/16	100	Dean (2007)
0/16	99/90	0/16	99/90	Irene (2011)

مدل‌هایی بر اساس داده‌های امواج دو طوفان مشهور Dean (2007) و Irene (2011) طی سه سال، برای پیش‌بینی وقوع امواج غیرعادی با استفاده از متغیرهای هواشناسی تهیه شد. متغیرهای هواشناسی مورد استفاده در مدل‌سازی‌ها عبارتند از: جهت باد، سرعت باد، ارتفاع موج مشخصه، فشار سطح دریا، دمای هوا و دمای سطح دریا. برای آموزش مدل‌ها، نواحی از طیف انرژی تبدیل مویک پیوسته که در آن‌ها انرژی بیشینه بود، به عنوان کاندیدای وقوع امواج غیرعادی در نظر گرفته شد. بدین ترتیب کلاس

7- نتیجه‌گیری

امواج واقعی دریا پدیده‌ای به شدت غیرخطی و تصادفی هستند که تاکنون رفتار آنها برای دانشمندان آن‌گونه که باید، شناخته نشده است. از طرفی ماهیت غیرخطی امواج غیرعادی بیشتر از سایر امواج است. به همین دلیل تلاش‌های انجام شده برای شناخت و پیش‌بینی وقوع این دسته از امواج توسط محققان نتیجه مطلوبی نداشته است. در این تحقیق با استفاده از مدل ترکیبی تبدیل مویک پیوسته و ماشین‌های بردار پشتیبان (WT-SVM)

waves”, *Coast. Eng. J.* 44 (4): 302–319.

Clauss, G. (2000). “Dramas of the sea: episodic wave sand their impact on offshore structures”, *Applied Ocean Research*, 24: 147–161, 2000.

Dean, R. (1990). “Abnormal waves: a possible explanation”, Torum, A., Gudmestad, O. (Eds.), *Water Wave Kinematics*. Kluwer, NewYork, 609–612.

Didenkova, I. (2011). “Shapes of freak waves in the coastal zone of the Baltic Sea (TallinBay)”, *Boreal Environ. Res.*, 16: 138–148.

Ewans, K., and Buchner, B. (2008). “Wavelet analysis of an extreme wave in a model basin”, In: *Proceedings of the 27th International Conference on Offshore Mechanics and Arctic Technology*, vol. 2: Structures, Safety and Reliability, Portugal. Paper No. OMAE 2008-57499: 455–464.

Haver, S. (2005). “Freak waves—a suggested definition and consequences for marine structures”, *Rogue Waves 2004*. IFREMER, Brest, France.

<http://www.ndbc.noaa.gov>

Huang, M. C. (2004). “Wave parameters and functions in wavelet analysis”, *Ocean Eng.* 31: 111–125.

Lavine, B.K., and Blank, T.R., (2009). 3.18 - Feed-forward neural networks, in *comprehensive chemometrics*, Edited by Steven D. Brown, Romá Tauler and Beata Walczak, Elsevier, Oxford, 2009, Pages 571-586, ISBN 9780444527011, <https://doi.org/10.1016/B978-044452701-1.00026-0>.

Liu, P., and Mori, N. (2000). “Characterizing freak waves with wavelet transform analysis”, *Rogue Waves 2000*, Brest, France, 151–156.

Massel, S. (2001). “Wavelet analysis for processing of ocean surface records”, *Ocean Eng.* 28: 957–987.

Scholkopf, B., and Smola, A.J. (2002). “*Learning with kernels*”. Cambridge, MA: The MIT Press.

Scholkopf, B., Burges, C., and Smola, A. (1998). “*Advances in kernel methods support vector learning*”. MIT Press.

Scholkopf, B., Mika, C.J., Burges, P, Knirsch, RR., and Muller, K.G. (1999). “Input space versus feature space in kernel-based methods”, *IEEE Trans Neural Network*, 10 (5):1000–1017.

Taylor, S.J. and Cristianini, N. (2004). “*Kernel methods for pattern analysis*”, Cambridge University

Tomita, H., and Kawamura, T. (2000). “Statistical analysis and inference from the in-situ data of the

امواج عادی از غیرعادی تمیز داده شد. از این داده‌ها به عنوان داده‌های هدف، و از متغیرهای هواشناسی به عنوان ورودی‌های ماشین بردار پشتیبان استفاده شد. مدل‌های ایجاد شده توانستند با دقت خوبی رابطه بین ورودی‌ها و خروجی‌ها را تخمین بزنند. بنابراین با رویکرد مورد استفاده در این تحقیق می‌توان در مواقعی که فقط اطلاعات هواشناسی در دسترس است، وقوع امواج غیرعادی را با دقت مناسبی تخمین زد.

7- فهرست علائم

d	انحراف ابر صفحه در طول بردار نرمال خود
b	پارامتر مقیاس
C	پارامتر تنظیم کننده
f	فرکانس
j	شمارنده
i	شمارنده
K	تابع کرنل
N	تعداد کلاس‌ها
N_x	تعداد داده‌های ورودی
t	زمان
w	بردار نرمال ابر صفحه جداکننده در فضای ویژگی
$x(t)$	سیگنال پیوسته
x_i	بردار داده‌های ورودی
x'_i	بردار داده‌های پیش‌بینی شده
y_i	کلاس داده‌های ورودی
*	بیانگر مزدوج مختلط
α	حدود انتگرال
δ_{mn}	تابع دلتای کرونگر
τ	پارامتر انتقال
$\phi(x_i)$	بردارهای تشابه در فضای ویژگی
$\psi_n(t)$	مجموعه‌ای از توابع شناخته شده
$\hat{\psi}_n$	توابع پایه متعامد برای ψ_n

8- منابع

Cherneva, Z., and Soares, C. (2014). “Time–frequency analysis of the sea state with the "Andrea" freak wave”, *Nat. Hazards Earth Syst. Sci.* 14: 3143–3150.

Chien, H., Kao, C. C., and Chuang, L. Z. H. (2002). “On the characteristics of observed coastal freak

Verlag; 2000, p. 131.

Veltcheva, A., and Guedes Soares, C. (2015). "Wavelet analysis of non-stationary sea waves during Hurricane Camille", *Ocean Technology* 95: 166–174.

Sea of Japan with reference to abnormal and/or freak waves", *Proceedings of the 10th ISOPE Conference*, Seattle, USA, 116–122.

Vapnik, VN. (2000). "*The nature of statistical learning theory*", 2nd ed. NY (USA): Springer-