

## Efficiency of machine learning models in estimation of local geoid height with GPS/Leveling measurements

Seyyed Reza Ghaffari-Razin\*<sup>1</sup> , Navid Hooshangi<sup>2</sup>

**Date of receive:**

2023.03.25

**Date of last review:**

2023.06.05

**Date of accept:**

2023.06.25

**Date of online publication:**

2023.08.11

**Extended Abstract**

**Introduction**

In geodesy, three levels are considered: the physical surface of the earth on which mapping measurements are made, the ellipsoidal reference surface (geometric datum) which is the basis of mathematical calculations, the geoid physical surface (physical datum) which is the basis for measuring heights. Satellite positioning systems measure the height of points relative to the ellipsoid surface. The geoid is one of the equipotential surfaces of the earth's gravity field, which approximates the mean sea level (MSL) by least squares. Geoid is very important in geodesy as a representative of the physical space or the space of observations made on the earth and also as the base level of elevations. The separation between the geoid and the geocentric reference ellipse is called geoid height (N). Although there is only one equipotential surface called geoid, various methods are used to determine it. These methods include: geometric method, geoid determination by satellite method, Gravimetric methods and geoid determination using GPS/leveling.

**Materials and Methods**

In this paper, the aim is to estimate the height of the local geoid using machine learning models. To do this, artificial neural network (ANN), adaptive neuro-fuzzy inference model (ANFIS), support vector regression (SVR) and general regression neural network (GRNN) models are used. The geodetic coordinates of 26 GPS stations in the north-west of Iran along with their orthometric height (H0) and normal height (h) were obtained from the national cartographic center of Iran. In all stations, the difference of orthometric height and normal height is considered as geoid height (N). Therefore, the geodetic longitude and latitude of the GPS stations are considered as the input of the machine learning models, and the corresponding geoid height was considered as the output. In order to test the results of machine learning models, two modes of 4 and 7 test stations are considered. Also, the output of the models is compared with the local geoid model IRG2016 presented by Saadat et al. for the Iranian region and also the global geoid model EGM2008. ... ▶ Page 100

**Keywords:**

Geoid,  
ANN,  
ANFIS,  
SVR,  
GRNN,  
Orthometric height

1- (\*corresponding author) Assistant professor, Department of Geoscience Engineering, Arak University of Technology, Arak, Iran, [mr.ghafari@arakut.ac.ir](mailto:mr.ghafari@arakut.ac.ir)

2- Assistant professor, Department of Geoscience Engineering, Arak University of Technology, Arak, Iran, [hooshangi@arakut.ac.ir](mailto:hooshangi@arakut.ac.ir)

### Results and Discussion

Due to the availability of a complete set of observations of GPS stations along with orthometric height obtained from leveling in the north-west region of Iran, the study and evaluation of the models proposed in the paper has been carried out in this region. Observations of 26 GPS stations of North-west of Iran were prepared from the national cartographic center (<https://www.ncc.gov.ir/>). Two modes are considered for training and testing of ANN, ANFIS, SVR and GRNN models. In the first case, the number of training stations is 22 and the number of test stations is 4. But in the second case, by increasing the number of test stations to 7 stations, the error evaluation of the models has been done. It should be noted that the distribution of training and test stations is completely random.

After the training step of machine learning models and choosing the optimal structure, the test step is performed in two different modes (4 and 7 stations). At this step, the value of the geoid height in the test stations is estimated and compared with the value obtained from the difference of orthometric height and normal height as a basis. Two statistical indices of relative error in percentage and RMSE in centimeters were calculated for all models and presented in Table (1) for the first case.

According to the results of Table (1) and comparing the relative error values of all models in the test stations, it shows that the ANFIS model was more accurate than other models. After ANFIS model, IRG2016 model has higher accuracy than ANN, SVR and GRNN models. It should be noted that the IRG2016 local model uses the observations of all Iranian plateau stations to model the local geoid, and therefore it is expected that this model will be more accurate in the study area than other models.

### Conclusion

The evaluations show that in the case of 22 training stations and 4 test stations, the RMSE of ANN, ANFIS, SVR, GRNN and IRG2016 models in the test step are 37.32, 19.83, 49.34, 53.82 and 29.65 cm, respectively. However, in the case of 19 training stations and 7 test stations, the error values of the models are 36.63, 58.31, 39.64, 41.29 and 24.68 cm, respectively. Comparison of RMSE shows that ANN model with less number of training stations provides higher accuracy than ANFIS, SVR and GRNN models. The results of this paper show that by using ANN and ANFIS models, geoid height can be estimated and used with high accuracy locally in civil and surveying applications.

**Table 1. Relative error (%) of ANN, ANFIS, SVR, GRNN and IRG2016 models in the test stations considered for the first case**

	TABZ	BZGN	SAGZ	AHAR
IRG2016	0.25	1.68	2.86	0.04
ANN	1.10	0.86	3.64	2.56
ANFIS	0.10	0.89	1.94	0.96
SVR	0.71	2.36	4.66	2.29
GRNN	0.74	1.83	5.92	2.06



## ارزیابی کارایی مدل‌های یادگیری ماشین

### در برآورد ارتفاع ژئوئید محلی با اندازه‌گیری‌های GPS/LEVELING

سیدرضا غفاری‌رزین<sup>۱</sup>، نوید هوشنگی<sup>۲</sup>

**چکیده**

در این مقاله کارایی مدل‌های شبکه عصبی مصنوعی (ANN)، سیستم استنتاج عصبی-فازی سازگار (ANFIS)، رگرسیون بردار پشتیبان (SVR) و مدل شبکه عصبی رگرسیون عمومی (GRNN) در تعیین ارتفاع ژئوئید محلی مورد ارزیابی قرار می‌گیرد. برای انجام این کار، مختصات ژئودتیکی ۲۶ ایستگاه از شبکه شمال غرب ایران که ارتفاع اورتومتريک ( $H^0$ ) آن‌ها نیز با ترازبایی درجه یک توسط سازمان نقشه‌برداری کشور (NCC) اندازه‌گیری شده، مورد استفاده قرار گرفته است. در این ایستگاه‌ها، تفاضل ارتفاع اورتومتريک از ارتفاع نرمال ( $h$ )، به‌عنوان ارتفاع ژئوئید (N) در نظر گرفته شده است. بنابراین ورودی مدل‌های ANN، ANFIS، SVR و GRNN مختصات طول و عرض ژئودتیکی ایستگاه‌ها بوده و خروجی متناظر با آن، ارتفاع ژئوئید است. آموزش مدل‌ها با استفاده از ۲۲ و ۱۹ ایستگاه انجام گرفته است. به‌عبارت دیگر تعداد ایستگاه‌های آموزش متغیر بوده تا بتوان آنالیز دقیق‌تری از دقت مدل‌ها را ارائه نمود. به‌منظور ارزیابی دقیق‌تر، نتایج با ژئوئید حاصل از مدل IRG2016 که توسط سازمان نقشه‌برداری کشور تولید شده، مقایسه می‌شوند. ارزیابی‌های انجام گرفته نشان می‌دهد که در حالت ۲۲ ایستگاه آموزش و ۴ ایستگاه آزمون، RMSE مدل‌های ANN، ANFIS، SVR و GRNN و IRG2016 در مرحله آزمون به ترتیب برابر با ۳۷/۳۲، ۱۹/۸۳، ۴۹/۳۴، ۵۳/۸۲ و ۲۹/۶۵ سانتی‌متر شده است. اما در حالت ۱۹ ایستگاه آموزش و ۷ ایستگاه آزمون، مقادیر خطای مدل‌ها به ترتیب برابر با ۳۶/۶۳، ۵۸/۳۱، ۳۹/۶۴، ۴۱/۲۹ و ۲۴/۶۸ سانتی‌متر به‌دست آمده است. مقایسه RMSE نشان می‌دهد که مدل ANN با تعداد ایستگاه‌های آموزش کمتر، دقت بالاتری نسبت به مدل‌های ANFIS و SVR ارائه می‌دهد. نتایج این مقاله نشان می‌دهد که با استفاده از مدل‌های ANN و ANFIS می‌توان ارتفاع ژئوئید را با دقت بالایی به‌صورت محلی برآورد کرده و مورد استفاده قرار داد.

\*\*\*\*\*

تاریخ دریافت:

۱۴۰۲/۰۱/۰۵

تاریخ آخرین بازنگری:

۱۴۰۲/۰۳/۱۵

تاریخ پذیرش:

۱۴۰۲/۰۴/۰۴

تاریخ انتشار:

۱۴۰۲/۰۵/۲۰

واژه‌های کلیدی:

ژئوئید،

مدل ANN

مدل ANFIS

مدل SVR

مدل GRNN

ارتفاع اورتومتريک

۱- استادیار گروه مهندسی نقشه‌برداری، دانشکده مهندسی علوم زمین، دانشگاه صنعتی اراک (نویسنده مسئول) mr.ghafari@arakut.ac.ir

۲- استادیار گروه مهندسی نقشه‌برداری، دانشکده مهندسی علوم زمین، دانشگاه صنعتی اراک hooshangi@arakut.ac.ir

## ۱- مقدمه

در تعیین ژئوئید با استفاده از روش هندسی که به نام روش استروژئودتیک<sup>۶</sup> نیز شناخته می‌شود به جای استفاده از اندازه بردار شتاب ثقل از راستای این بردار برای تعیین ژئوئید استفاده می‌شود. منظور از راستای بردار شتاب ثقل، همان زوایای انحراف قائم<sup>۷</sup> است که می‌توان از این زوایا به‌عنوان مشتقات سوئی ژئوئید در دو راستای نصف‌النهاری و مقطع قائم اولیه نام برد. از این رو تعیین ژئوئید با استفاده از این مشاهدات قابل قیاس با حل یک معادله دیفرانسیل با مشتقات جزئی مرتبه اول است (اردلان، ۱۹۹۹). تابع پتانسیل جاذبه زمین قابل بسط به هارمونیک‌های کروی است. ضرایب هارمونیک‌های با درجه و مرتبه پایین در این بسط را می‌توان با استفاده از آنالیز اغتشاشات مدار ماهواره‌های در حال گردش به دور زمین همانند ماهواره‌های ثقل‌سنجی<sup>۸</sup> تعیین نمود (صفری و همکاران، ۲۰۰۵). پر واضح است که با توجه به پوشش یکنواخت و تراکم بالای داده‌های ماهواره‌ای، ضرایب طول موج‌های بلند پتانسیل جاذبه زمین که از این مشاهدات محاسبه می‌شوند بسیار دقیق‌تر از ضرایب محاسبه شده با استفاده از مشاهدات زمینی باشند. از سوی دیگر به دلیل فاصله زیاد ماهواره تا سطح زمین، مشاهدات ماهواره‌ای فاقد جزئیات میدان جاذبه یا طول موج‌های کوتاه آن هستند. در نتیجه ژئوئید حاصل از این روش با وجود همگن بودن و داشتن کیفیت یکنواخت در همه جای زمین، به دلیل نداشتن طول موج‌های کوتاه بسیار هموار است (نجفی، ۲۰۰۴). روش جاذبی تعیین ژئوئید مبتنی بر به‌کارگیری مشاهدات اندازه بردار شتاب ثقل زمین است (ونچیک و کراکوسکی، ۱۹۸۶). در این روش با حل یک مسئله مقدار مرزی<sup>۹</sup> با استفاده از معادله دیفرانسیل لاپلاس برای تابع پتانسیل جاذبه در خارج زمین، می‌توان ژئوئید را تعیین نمود. روش استوکس-هلمرت یک از روش‌های جاذبی تعیین ژئوئید است. بایستی اشاره شود که در این

آگاهی از شکل و اندازه زمین در اهداف ژئوفیزیک و علوم کاربردی به‌منظور تعیین موقعیت دقیق عوارض در سطح زمین از اهمیت حیاتی برخوردار است. در ژئودزی<sup>۱</sup>، سه سطح مطرح است: سطح فیزیکی زمین که اندازه‌گیری‌های نقشه‌برداری بر روی آن انجام می‌گیرد، سطح بیضوی مرجع (دیتوم هندسی)<sup>۲</sup> که مبنای محاسبات ریاضی است، سطح فیزیکی ژئوئید<sup>۳</sup> (دیتوم فیزیکی) که مبنای سنجش ارتفاعات است. سامانه‌های تعیین موقعیت ماهواره‌ای، ارتفاع نقاط را نسبت به سطح بیضوی اندازه‌گیری می‌کند. اما در بسیاری از پروژه‌های نقشه‌برداری که ارتفاع و شیب مهم است، ارتفاع از سطح بیضوی کارایی ندارد، زیرا یک سطح ریاضی است و در پروژه‌های وسیعی که ارتفاع و شیب در آن اهمیت دارد، مناسب نیست. راه‌حل درست برای این مشکل، تعریف یک سطح فیزیکی است که نزدیک‌ترین شکل را به زمین داشته باشد. این سطح همان دیتوم فیزیکی یا ژئوئید است. ژئوئید یکی از سطوح هم‌پتانسیل میدان ثقل زمین است که سطح متوسط آب‌های آزاد<sup>۴</sup> (MSL) را با نرم‌کمترین مربعات تقریب می‌نماید. ژئوئید به‌عنوان نماینده فضای فیزیکی یا فضای مشاهدات انجام شده روی زمین و همچنین به‌عنوان سطح مبنای ارتفاعات، دارای اهمیت بسیار زیادی در ژئودزی است. جدایی بین ژئوئید و بیضوی فرانس ژئوسنتریک را ارتفاع ژئوئید<sup>۵</sup> (N) می‌نامند (هافمن و موریتز، ۲۰۰۶). با وجود اینکه تنها یک سطح هم‌پتانسیل به نام ژئوئید وجود دارد، ولی روش‌های گوناگونی برای تعیین آن به کار برده می‌شود. این روش‌ها عبارتند از: روش هندسی تعیین ژئوئید (ونچیک و کراکوسکی، ۱۹۸۶)، تعیین ژئوئید به روش ماهواره‌ای (هافمن و موریتز، ۲۰۰۶). روش‌های جاذبی تعیین ژئوئید (نجفی علمداری، ۲۰۰۴؛ وبر و زمردیان، ۱۹۸۸) و تعیین ژئوئید با استفاده از GPS/leveling (کنیز، ۲۰۱۶).

6- Astrogeodetic method  
7- Deflection of vertical  
8- Gravitational satellites  
9- Boundary value problem

1- Geodesy  
2- Geometrical datum  
3- Geoid  
4- Mean Sea Level  
5- Geoid height

## فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (۳۳)

ارزیابی کارائی مدل‌های یادگیری ماشین در برآورد ارتفاع ژئوئید محلی با اندازه‌گیری‌های ... / ۱۰۳

تعیین ارتفاع ژئوئید دقیق محاسبه کردند. پس از آن حاتم (۲۰۱۰) مدل ژئوئید محلی دقیق ایران بنام IRGeoid10 را براساس روش استوکس، تکنیک حذف و اضافه و روش تبدیل فوریه سریع یک بعدی<sup>۳</sup>، ارائه نمود. او به دقتی در حدود ۲۶ سانتی‌متر بر روی ۸۱۹ ایستگاه کنترل GNSS/Leveling موجود در کشور دست یافت. در نهایت سعادت و همکاران (۲۰۱۸) ژئوئید محلی دقیق ایران به نام IRG2016 را براساس تلفیق داده‌های زمینی و گرادیومتری ماهواره‌ای بر مبنای توابع پایه شعاعی<sup>۴</sup> (RBF) ارائه کردند. آن‌ها به دقتی در حدود ۲۳ سانتی‌متر بر روی ۱۲۸۸ ایستگاه کنترل GNSS/Leveling پراکنده شده در سراسر ایران دست یافتند.

امروز با گسترش سامانه‌های تعیین موقعیت ماهواره‌ای همانند GPS، مشاهده مستقیم کمیت ارتفاع ژئوئید با استفاده از سامانه‌های مذکور و مشاهدات ترازیبی دقیق امکان‌پذیر شده است (سبری و همکاران، ۲۰۱۹؛ فروغی و همکاران، ۲۰۲۳). با استفاده از مشاهدات ترازیبی دقیق، ارتفاع اورتومتريک<sup>۵</sup> نقاط شبکه حاصل می‌شود. حال چنانچه با استفاده از سامانه تعیین موقعیت جهانی، در این نقاط ارتفاع ژئودتیکی<sup>۶</sup> نیز اندازه‌گیری شود، از اختلاف این دو کمیت یعنی ارتفاع ژئودتیکی و ارتفاع اورتومتريکی می‌توان ارتفاع ژئوئید را در نقاط مورد نظر با دقت بسیار بالایی به دست آورد. نقاطی که دارای هر دو نوع ارتفاع ژئودتیکی و اورتومتريکی هستند به نقاط GPS/Leveling معروف هستند. واضح است که دقت ارتفاع ژئوئید تعیین شده از این روش به‌طور مستقیم به دقت ارتفاعات به دست آمده از دو روش مذکور بستگی دارد. با وجود اینکه ژئوئید به دست آمده از این روش دارای تمامی طول موج‌های مختلف این کمیت است اما به دلیل محدودیت موجود در کار ترازیبی دقیق، دارای پخش یکنواختی در همه نقاط سطح زمین نیست.

روش علاوه بر مشاهدات شتاب ثقل، اطلاعات دیگری از پوسته زمین همانند مدل رقومی و چگالی پوسته زمین مورد نیاز است. با توجه به اینکه مشاهدات شتاب ثقل در نقاط محدودی از سطح زمین وجود دارد، در نتیجه ژئوئید حاصل از این روش دارای دقت و کیفیت یکنواختی نبوده و همچنین دارای بایاس طول موج‌های بلند است (نهادنچی و سلطانپور، ۲۰۰۵).

تعیین مدل ژئوئید دقیق در ایران موضوع جدیدی نیست. تلاش‌های بسیاری توسط محققین و پژوهشگران مختلف در سال‌های گذشته در این خصوص انجام شده است. نخستین تحقیقات در این زمینه مربوط به ۳۰ سال قبل بوده، جایی که وِبر و زمردیان (۱۹۸۸) روش انتگرالی را برای بهبود مدل ژئوپتانسیلی متناسب با داده‌های آنومالی ثقل<sup>۱</sup> زمینی برای تعیین ژئوئید محلی در ایران به کار بردند. آن‌ها به دقت اولیه‌ای در حدود  $\pm 1/4$  متر در مقایسه با مشاهدات داپلر و ترازیبی دقیق در منطقه مورد مطالعه دست یافتند. پس از آن هامش و زمردیان (۱۹۹۲) با استفاده از مدل رقومی ارتفاعی برگرفته از نقشه‌های ۱:۲۵۰,۰۰۰ اثر تصحیحات توپوگرافی<sup>۲</sup> را بر روی داده‌های شتاب‌ثقل موجود در بانک اطلاعاتی BGI محاسبه نموده و با استفاده از تکنیک حذف، محاسبه و اضافه، یک مدل ژئوئید محلی برای ایران محاسبه کردند. سپس اردلان و همکاران (۲۰۰۲) مدل ژئوئید دقیق ایران را براساس روش پتانسیلی (اردلان، ۱۹۹۹) بر مبنای فرمول برنز تعمیم یافته به دست آوردند. همچنین نجفی (۲۰۰۴) و نهادنچی و سلطانپور (۲۰۰۵) روش استوکس-هلمرت را برای تعیین ارتفاع ژئوئید دقیق در ایران به کار بردند. پس از آن صفری و همکاران (۲۰۰۵) مدل ژئوئید دقیق ایران را با استفاده از تقریب بیضوی فرمول برنز و حل معکوس انتگرال آبل-پوآسن تعیین کردند. سپس کیامهر (۲۰۰۶) مدل ژئوئید دقیق ایران را با اصلاح کمترین مربعات کرنل استوکس معرفی نمود. همچنین صدیقی و همکاران (۲۰۰۸) اثر تغییرات چگالی سطحی را بر روی

3- One-dimensional fast Fourier transform

4- Radial Basis Functions

5- Orthometric height

6- Geodetic height

1- Gravity anomaly

2- Topographic corrections

و همکاران (۲۰۰۷) از مدل ماشین بردار پشتیبان<sup>۶</sup> (SVM) با کرنل‌های موجکی برای مدل‌سازی ارتفاع ژئوئید با استفاده از دیتاهای GPS/Leveling پرداختند. آموزش مدل SVM در راستای دستیابی به بالاترین دقت و با در نظر گرفتن عدم به وجود آمدن مسئله فیت بیش از حد<sup>۷</sup> انجام شد. نتایج حاصل از این تحقیق نشان‌دهنده RMSE<sup>۸</sup> سه سانتی‌متری برای ارتفاع ژئوئید حاصل از مدل SVM بود.

## ۲- منطقه مورد مطالعه و مشاهدات مورد استفاده

با توجه به در اختیار بودن مجموعه کاملی از مشاهدات ایستگاه‌های GPS به همراه ارتفاع اورتومتريک حاصل از ترازیابی درجه یک در منطقه شمال‌غرب ایران، مطالعه و ارزیابی مدل‌های پیشنهادی مقاله در این ناحیه انجام گرفته است. مشاهدات ۲۶ ایستگاه GPS شبکه شمال‌غرب ایران از سازمان نقشه‌برداری کشور تهیه شده است (<https://www.ncc.gov.ir>).

دو حالت به‌منظور آموزش و آزمون مدل‌های ANN، ANFIS، SVR و GRNN در نظر گرفته شده است. در حالت اول تعداد ایستگاه‌های آموزش ۲۲ و تعداد ایستگاه‌های آزمون ۴ است. اما در حالت دوم با افزایش تعداد ایستگاه‌های آزمون به ۷ ایستگاه، ارزیابی خطای مدل‌ها انجام گرفته است. به عبارت دیگر در سناریوی دوم، تعداد ایستگاه‌های مورد استفاده برای آموزش کمتر شده تا ارزیابی دقیق‌تری از خطای مدل‌های یادگیری ماشین صورت پذیرد. بایستی اشاره شود که چگونگی پراکندگی ایستگاه‌های آموزش و آزمون به‌صورت کاملاً تصادفی است. همچنین مشاهدات ایستگاه‌های آزمون در مرحله آموزش کنار گذاشته شده است.

نگاره (۱) چگونگی پراکندگی ایستگاه‌های مورد استفاده در این مقاله را نمایش می‌دهد.

هدف اصلی این مقاله استفاده از قابلیت‌های مدل‌های یادگیری ماشین<sup>۱</sup> (ML) برای مدل‌سازی محلی ارتفاع ژئوئید با استفاده از مشاهدات ایستگاه‌های GPS/Leveling است. برای این منظور از چهار مدل شبکه عصبی مصنوعی<sup>۲</sup> (ANN)، سیستم استنتاج عصبی-فازی سازگار<sup>۳</sup> (ANFIS)، مدل رگرسیون بردار پشتیبان<sup>۴</sup> (SVR) و مدل شبکه عصبی رگرسیون عمومی<sup>۵</sup> (GRNN) برای مدل‌سازی محلی ارتفاع ژئوئید استفاده می‌شود. به عبارت دیگر ورودی چهار مدل، مختصات ژئودتیکی ایستگاه بوده و خروجی مطلوب، ارتفاع ژئوئید خواهد بود.

نوآوری اصلی این مقاله در مدل‌سازی ارتفاع ژئوئید محلی با استفاده از مدل‌های SVR و GRNN برای اولین بار در ایران است. همچنین استفاده از دو سناریوی مختلف با تعداد ایستگاه‌های آموزشی متفاوت از قابلیت‌های بارز این تحقیق است. تاکنون تحقیقات معدودی در زمینه استفاده از قابلیت مدل‌های یادگیری ماشین برای تعیین ارتفاع ژئوئید در ایران و جهان انجام گرفته است. حضرائی و همکاران (۱۳۹۳) به ارزیابی روش‌های نوین محاسباتی مبتنی بر یادگیری از جمله مدل ANN و ANFIS در مقایسه با روش استاندارد معادلات رگرسیون چندجمله‌ای چند متغیره، در مدل‌سازی ژئوئید GPS/Leveling در منطقه شاهین‌شهر اصفهان پرداختند. نتایج این تحقیق نشان داد که مدل ANFIS از دقت بالاتری برخوردار است. کاوزوگلو و ساکا (۲۰۰۵) از دیتاهای GPS/leveling برای مدل‌سازی ارتفاع ژئوئید در منطقه استامبول ترکیه استفاده کردند. مدل‌های ANN، چندجمله‌ای، کالوکیشن کمترین مربعات و المان محدود به‌منظور مقایسه و ارزیابی انتخاب شدند. نتایج حاصل از این تحقیق نشان داد که مدل ANN دارای قابلیت بالایی در مدل‌سازی ارتفاع ژئوئید در این منطقه است و برای کارهای اجرایی از دقت قابل قبولی برخوردار است. زالتنیک

1- Machine Learning models

2- Artificial Neural Network

3- Adaptive Neuro-Fuzzy Inference System

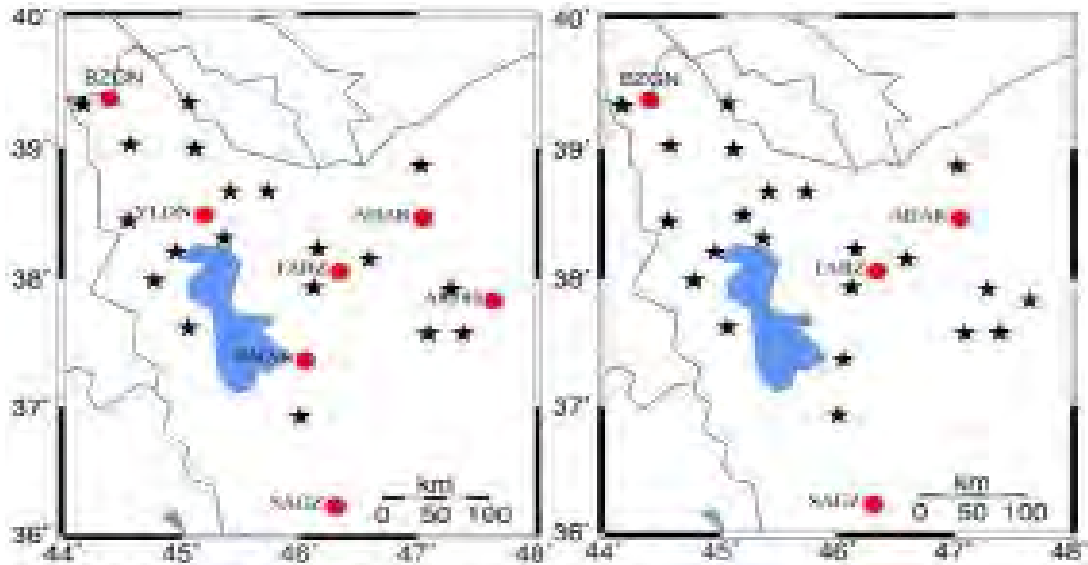
4- Support Vector Regression model

5- General Regression Neural Network model

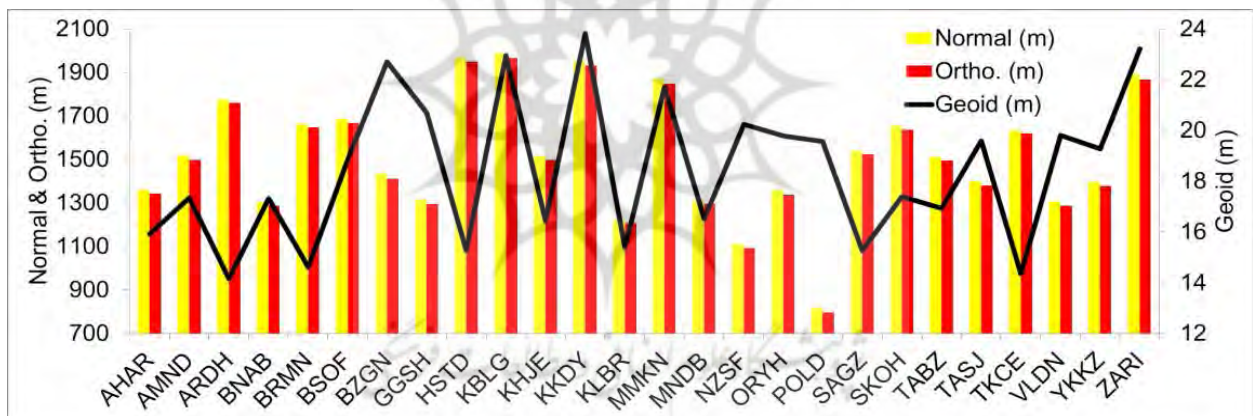
6- Support Vector Machine

7- Overfitting problem

8- Root Mean Square Error



نگاره ۱: چگونگی پراکنندگی ایستگاه‌های آموزش (ستاره‌های سیاه رنگ) و ایستگاه‌های کنترل (دایره‌های قرمز رنگ) در حالت اول (شکل راست) و حالت دوم (شکل چپ)



نگاره ۲: چگونگی تغییرات ارتفاع نرمال (ستون‌های زرد رنگ)، ارتفاع اورتومتریک (ستون‌های قرمز رنگ) و ارتفاع ژئوئید (منحنی سیاه رنگ) در کلیه ایستگاه‌های مورد استفاده در این تحقیق

توجه به اینکه ارتفاع ژئوئید، حاصل تفاضل ارتفاع نرمال و ارتفاع اورتومتریک است، در نتیجه مطابق با قانون انتشار خطاها، ارتفاع ژئوئید حاصل از این تفاضل، دارای خطایی در حدود ۳۵/۳۴ میلی‌متر (۳/۵۳ سانتی‌متر) است. ارتفاع نرمال یک کمیت هندسی است و نسبت به دیتوم WGS84 اندازه‌گیری می‌شود. بنابراین اگر دیتوم تغییر نکند و بنچ‌مارک‌های موجود در منطقه از لحاظ ارتفاعی تغییری نداشته باشند، در نتیجه این ارتفاع در طول زمان ثابت

ارتفاع کلیه ایستگاه‌های مورد استفاده از ۱۲۸۰ تا ۱۹۵۲ متر متغیر است. همچنین برای غلبه بر خطای چندمسیری، در این ایستگاه‌ها از آنتن چوک رینگ با زاویه ارتفاعی ۱۵ درجه استفاده شده است. نگاره (۲) نشان‌دهنده سه ارتفاع نرمال، اورتومتریک و ارتفاع ژئوئید برای کلیه ایستگاه‌های مورد استفاده در این تحقیق است.

بایستی اشاره شود که دقت ارتفاع اورتومتریک و نرمال به ترتیب برابر با ۸/۵۵ و ۳۴/۲۸ میلی‌متر است. بنابراین با

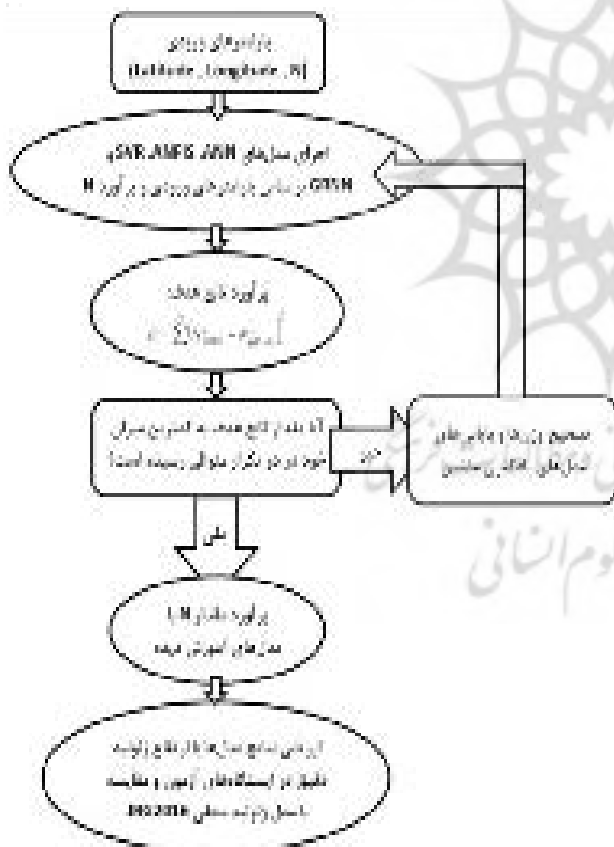


محاسبه می شوند:

$$Re. (\%) = \frac{|N_{model} - N_{(H^0-h)}|}{VTEC_{GPS}} \times 100 \quad \text{رابطه (۳)}$$

$$RMSE = \sqrt{\frac{1}{Q} \sum_{q=1}^Q (N_{model}^q - N_{(H^0-h)}^q)^2} \quad \text{رابطه (۴)}$$

کمینه مقادیر خطای نسبی و RMSE صفر است. هر چقدر مقادیر این دو پارامتر به صفر نزدیکتر باشند، نشان دهنده دقت بالاتر مدل مورد نظر است. فلوجارت چگونگی انجام آموزش و آزمون مدل های یادگیری ماشین در نگاره (۳) نمایش داده شده است.



نگاره ۳: فلوجارت چگونگی استفاده از مدل های ANN، ANFIS، SVR و GRNN برای برآورد ارتفاع ژئوئید و ارزیابی دقت آن

خواهد بود. برای ارتفاع اورتومتريک که یک کمیت فیزیکی است و نسبت به دیتوم فیزیکی یعنی ژئوئید اندازه گیری می شود همچنین حالتی وجود دارد. اگر دیتوم ثابت باشد و ایستگاه ها در طول زمان جابجایی نداشته باشند، بنابراین ارتفاع اورتومتريک یک نقطه هم ثابت خواهد بود. با فرض ثابت بودن دو ارتفاع فوق در طول زمان، از تفاضل این دو ارتفاع (اورتومتريک و نرمال)، ارتفاع ژئوئید تمامی ایستگاه ها در این مقاله محاسبه شده است. ارتفاع ژئوئید به دست آمده، در حقیقت مرجع اصلی برای مقایسه ها در ایستگاه های آزمون است.

ورودی مدل های یادگیری ماشین طول و عرض جغرافیایی ایستگاه های آموزش بوده و خروجی متناظر، ارتفاع ژئوئید ایستگاه خواهد بود. بنابراین رابطه ریاضی (۱) را می توان در نظر گرفت:

$$N_{ANN, ANFIS, SVR, GRNN} = f(latitude, longitude) \quad \text{رابطه (۱)}$$

در رابطه (۱) بیانگر ارتفاع ژئوئید حاصل از مدل های یادگیری ماشین، latitude و longitude به ترتیب نشان دهنده عرض و طول جغرافیایی ایستگاه GPS است. تابع هزینه در نظر گرفته برای چهار مدل ANN، ANFIS، SVR و GRNN به صورت رابطه (۲) است:

$$E = \sum_{q=1}^Q (N_{model}^q - N_{(H^0-h)}^q)^2 \quad \text{رابطه (۲)}$$

در رابطه (۲) بیانگر ارتفاع ژئوئید حاصل از مدل های یادگیری ماشین،  $N_{(H^0-h)}^q$  نشان دهنده ارتفاع ژئوئید در ایستگاه های آزمون و Q تعداد نقاط آزمون است. ارزیابی نتایج مدل های ANN، ANFIS، SVR و GRNN در پیش بینی ارتفاع ژئوئید با استفاده از مقادیر ارتفاع ژئوئید در ایستگاه های آزمون (تفاضل ارتفاع اورتومتريک از ارتفاع نرمال) به عنوان مشاهده مرجع انجام می گیرد. پارامترهای آماری خطای نسبی و RMSE برای بررسی دقت و صحت مدل ها استفاده می شوند. این پارامترها با روابط (۳ و ۴)



## فصلنامه علمی - پژوهشی اطلاعات جغرافیایی (۳۳)

ارزیابی کارائی مدل‌های یادگیری ماشین در برآورد ارتفاع ژئوئید محلی با اندازه‌گیری‌های ... / ۱۰۷

مناسب نیاز است. یکی از مناسب‌ترین توابع فعال‌سازی تابع سیگموئید است که به صورت تابع ریاضی (۶) تعریف می‌شود:

$$f(z) = \frac{1}{1+e^{-z}} \quad \text{رابطه (۶)}$$

که در رابطه (۶)  $z$  اطلاعات ورودی نرون و  $f(z) \in [0,1]$  هستند. برای آموزش شبکه عصبی، الگوریتم‌های آموزش متفاوتی وجود دارد. در مقاله حاضر، از الگوریتم آموزش پس‌انتشار خطا (BP) استفاده می‌شود (مارس و همکاران، ۱۹۹۶). در این الگوریتم خروجی شبکه عصبی با خروجی مطلوب مقایسه شده و براساس تابع هدف (رابطه (۲))، مقدار خطا محاسبه می‌شود. اگر مقدار خطا در حد دقت مطلوب نباشد، براساس مشتق تابع هدف، وزن‌های مربوط به نرون‌ها تصحیح شده و مجدداً خروجی محاسبه می‌شود. این مراحل تا مادامی که مقدار خطا (تابع هدف) به کمترین میزان خود برسد، ادامه می‌یابد.

مطابق با فلوجارت ارائه شده در نگاره (۳) فرآیند اعتبارسنجی مدل‌ها در ایستگاه‌های آزمون انجام گرفته است. به این صورت که ارتفاع ژئوئید دقیق حاصل از تفاضل ارتفاع اورتومتريک و ارتفاع نرمال، در این ایستگاه‌ها مشخص است. پس از مرحله آموزش مدل‌های یادگیری ماشین، مختصات ژئودتیکی ایستگاه‌های آزمون به مدل‌ها معرفی شده و ارتفاع ژئوئید با مدل‌های یادگیری ماشین برآورد می‌شود. این ارتفاع با ارتفاع ژئوئید دقیق (تفاضل ارتفاع اورتومتريک و ارتفاع نرمال) مقایسه شده و خطای مدل‌ها در ایستگاه‌های آزمون محاسبه شده است.

### ۳- روش تحقیق

در این بخش در مورد تئوری ریاضی مدل‌های ANN، SVR، ANFIS، GRNN و مدل محلی ژئوئید IRG2016 توضیحات مفصلی ارائه می‌شود.

### ۳-۱- شبکه عصبی مصنوعی

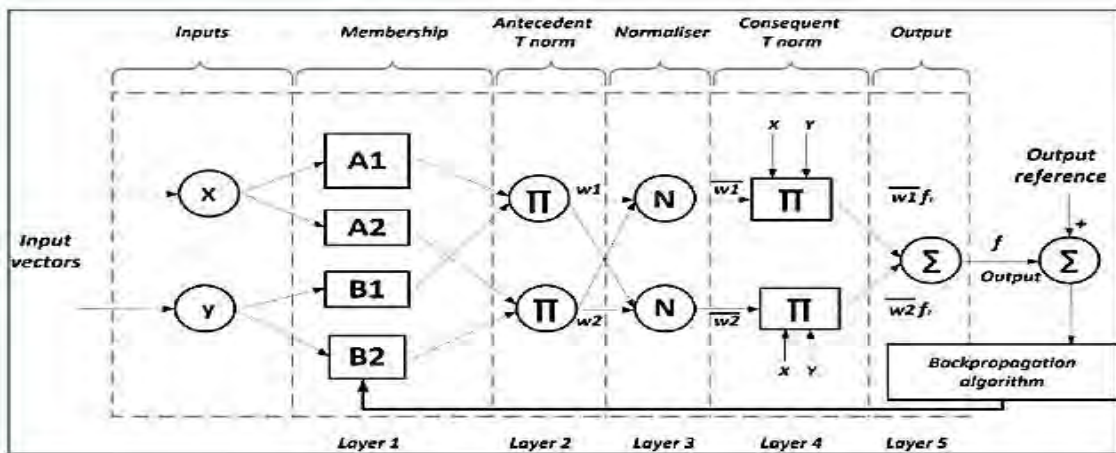
در این مقاله برای مدل‌سازی ارتفاع ژئوئید از یک شبکه عصبی پرسپترون سه لایه استفاده می‌شود. حسن این کار در تسریع عملیات برآورد پارامترها و نوع یادگیری و آموزش شبکه است. شبکه عصبی پرسپترون سه لایه مورد استفاده دارای یک لایه ورودی با  $N$  مشاهده، یک لایه پنهان و یک لایه خروجی با  $n$  پارامتر است. مدل مورد استفاده برای محاسبه پارامترهای لایه خروجی به صورت رابطه (۵) است (هایکین، ۱۹۹۴):

$$y = f\left(\sum_{j=1}^q W_j f\left(\sum_{l=1}^N w_{j,l} x_l + w_{j,0}\right) + W_0\right) \quad \text{رابطه (۵)}$$

در رابطه (۵)  $W_j$  وزن مابین زمین نرون لایه پنهان و خروجی،  $w_{j,l}$  وزن مابین زمین نرون ورودی و زمین نرون پنهان،  $x_l$  بیانگر زمین پارامتر ورودی،  $w_{j,0}$  بایاس مربوط به نرون‌های مابین لایه‌های ورودی و پنهان و  $W_0$  بایاس مربوط به نرون‌های لایه‌های پنهان و خروجی هستند. برای ارتباط میان لایه پنهان و خروجی به یک تابع فعال‌سازی

### ۳-۲- سیستم استنتاج عصبی-فازی سازگار

مدل‌های عصبی-فازی که به واسطه جنگ در سال ۱۹۹۳ گسترش یافت، برای تسهیل فرآیند یادگیری و انطباق، منطق فازی را با شبکه‌های عصبی مصنوعی ترکیب می‌کند. در واقع در مدل‌های عصبی-فازی برای حل مشکل شناسایی پارامترهای سیستم استنتاج فازی (FIS)، از یک شبکه تطبیقی که حالت عمومی شبکه عصبی پیشرو چندلایه است، استفاده می‌شود. همانند سیستم‌های فازی، ساختار ANFIS نیز از دو بخش تشکیل شده است. بخش نخست مقدم و بخش دوم تالی (استنتاج - نتیجه) نامیده می‌شود که این دو بخش توسط قواعد فازی در فرم یک شبکه به یکدیگر متصل می‌شوند. نگاره (۴) ساختار یک شبکه ANFIS را در پنج لایه نمایش می‌دهد، که در آن لایه اول عمل فازی‌سازی را انجام می‌دهد، لایه دوم عمل T-norm فازی را برای بخش مقدم قواعد فازی انجام می‌دهد، لایه



نگاره ۴: ساختار شبکه ANFIS دارای دو ورودی

$$f_{jk} = q_{0,jk} + q_{1,jk} \cdot x_1 + q_{2,jk} \cdot x_2 \quad \text{رابطه (۹)}$$

که در روابط فوق  $m$  نشان‌دهنده تعداد توابع تعلق برای هر متغیر ورودی بوده و  $(q_0, q_1, q_2)$  اعداد حقیقی متناظر با وزن‌های خطی در بخش تالی سیستم ANFIS هستند. خروجی نهایی  $y$  شبکه ANFIS به صورت رابطه (۱۰) قابل محاسبه است:

$$y = \frac{\sum_{j=1}^M \sum_{k=1}^N f_{jk} \mu_{M^j_1}(x_1) \mu_{M^k_2}(x_2)}{\sum_{j=1}^M \sum_{k=1}^N \mu_{M^j_1}(x_1) \mu_{M^k_2}(x_2)} = \sum_{j=1}^M \sum_{k=1}^N f_{jk} \bar{w}_{jk} \quad \text{رابطه (۱۰)}$$

به منظور مدل‌سازی سیستم‌های غیرخطی پیچیده، مدل ANFIS فضای ورودی را به بخش‌های مختلفی افزایش می‌نماید. به عبارتی فضای ورودی به نواحی محلی فراوانی تقسیم می‌شود. شبکه ANFIS از توابع تعلق فازی به منظور تقسیم هر بعد ورودی استفاده می‌نماید. این توابع تعلق با یکدیگر همپوشانی دارند، به عبارتی یک ورودی منفرد سبب فعال شدن همزمان حداقل دو تابع تعلق خواهد شد. قابلیت شبکه ANFIS به تعداد توابع تعلق در نظر گرفته شده برای هر بعد ورودی وابسته است. معمولاً توابع تعلق استفاده شده توابع زنگی شکل گوسی با میزان بیشینه برابر با یک و کمینه برابر با صفر هستند که به فرم رابطه (۱۱)

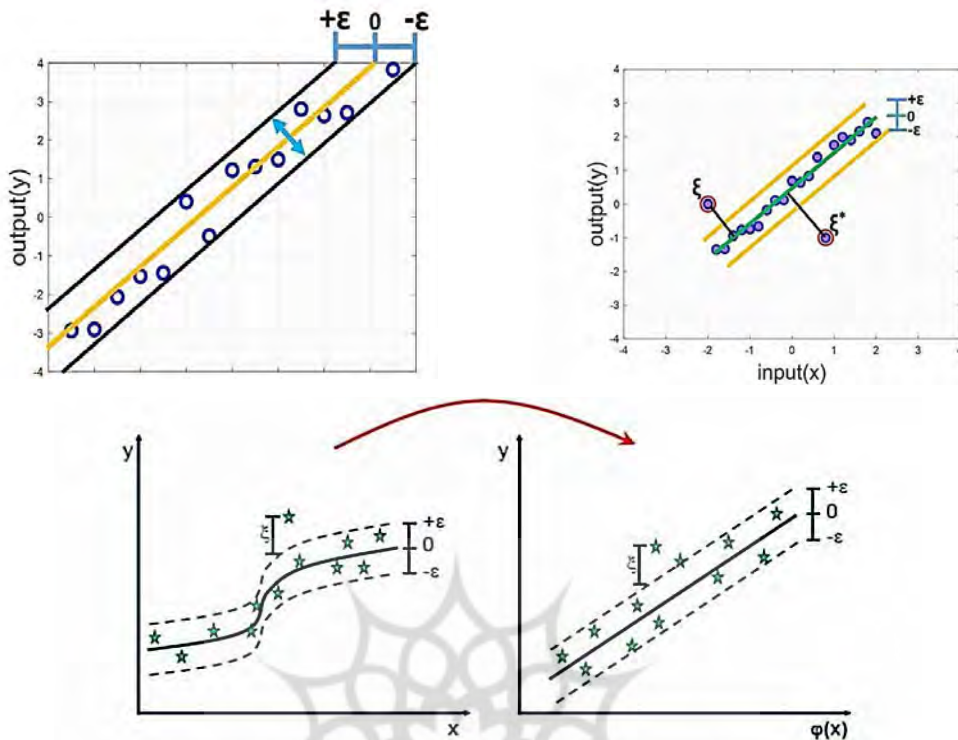
سوم به منظور نرمال‌سازی به کار می‌رود، لایه چهارم بخش تالی قواعد فازی را ایجاد می‌نماید و در نهایت لایه پنجم خروجی نهایی سیستم را محاسبه می‌کند.

سیستم استنتاج عصبی-فازی سازگار شبکه‌ای پنج لایه‌ای، متشکل از گره‌ها و کمان اتصال دهنده گره است. لایه اول، داده‌های ورودی با درجه عضویت است که توسط کاربر مشخص می‌شود. کلیه عملیات مدل‌سازی در لایه‌های دوم تا چهارم انجام می‌گیرد. لایه آخر، خروجی شبکه است که هدف آن حداقل نمودن اختلاف خروجی به دست آمده از شبکه و خروجی واقعی است. ساختار مناسب ANFIS متناسب با داده‌های ورودی، درجه عضویت، قوانین و توابع درجه عضویت خروجی انتخاب می‌شود. در مرحله آموزش با اصلاح پارامترهای درجه عضویت بر اساس میزان خطای قابل قبول، مقادیر ورودی به مقادیر واقعی نزدیک‌تر می‌شوند. روابط پیشخور برای ساختار شبکه نمایش داده شده در نگاره (۴) به شرح روابط (۷، ۸ و ۹) هستند:

$$w_{jk} = \mu_{M^j_1}(x_1) \mu_{M^k_2}(x_2) \quad \text{رابطه (۷)}$$

$$j, k = 1, 2, \dots, m$$

$$\bar{w}_{jk} = \frac{w_{jk}}{\sum_{i=1}^m \sum_{f=1}^m w_{if2}} \quad j, k = 1, 2, \dots, m \quad \text{رابطه (۸)}$$



نگاره ۵: رگرسیون خطی با مدل SVR (ردیف بالا-راست)، خطی با داده‌های پرت (ردیف بالا-چپ) و غیرخطی (ردیف پایین)

همکاران، ۲۰۱۲). نگاره (۵) حالت‌های مختلف مدل SVR را برای رگرسیون داده‌ها نشان می‌دهد.

در حالت رگرسیون خطی، مدل SVR دنبال رابطه خطی است که از وسط خروجی‌ها عبور کند و حاشیه‌ای ایجاد کند که در آن خروجی همه نمونه‌ها در داخل حاشیه قرار گیرند. با چنین رویکردی می‌توان بهترین رابطه خطی ممکن بین ورودی و خروجی را به دست آورد. در حالت رگرسیون غیرخطی، مدل SVR نیز از رویکرد مدل SVM استفاده کرده و به تابع هزینه یک متغیر مجازی<sup>۲</sup> اضافه می‌کند تا بتواند مشکل داده‌های پرت را حل کند. به منظور تعمیم مدل SVR برای حل مسائل غیرخطی نیز از رویکردی مشابه با حالت وجود نمونه‌های پرت در داده‌های ورودی استفاده می‌شود. یعنی در ابتدا داده‌ها با کمک یک نگاشت غیرخطی به فضای خطی تصویر می‌شوند و سپس در فضای جدید، رابطه خطی بین داده‌ها در فضای ویژگی و خروجی به دست می‌آید. این

قابل تعریف هستند:

$$\mu_{M_i}(x) = \exp \left\{ - \left( \frac{x - \bar{x}_i}{\sigma_i} \right)^2 \right\} \quad \text{رابطه (۱۱)}$$

که در رابطه (۱۱)  $(\bar{x}_i, \sigma_i)$  پارامترهای توابع تعلق هستند که بر روی شکل آن تأثیر می‌گذارند.

### ۳-۳- رگرسیون بردار پشتیبان

روش رگرسیون بردار پشتیبان یک روش یادگیری ماشین مبتنی بر هسته برای طبقه‌بندی و رگرسیون است که در آن ریسک عدم طبقه‌بندی صحیح را کمینه می‌کنند. ساختار یک شبکه SVR، اشتراکات زیادی با شبکه عصبی پرسپترون چند لایه دارد و تفاوت اصلی آن عملاً در شیوه یادگیری است. به طور کلی این روش به دو حالت خطی و غیرخطی تقسیم می‌شود. همچنین می‌توان در هر دو حالت اثر داده‌های پرت<sup>۱</sup> را در مدل‌سازی در نظر گرفت (بگانه و

2- Slack variable

1- Outliers

کارایی دارد. از نظر ساختاری، شبکه عصبی رگرسیون عمومی دارای ساختاری مشابه با شبکه عصبی تابع شعاعی با یک لایه اضافه از نوع خطی، قبل از خروجی است. نقش این لایه محاسبه یک رگرسیون از داده‌های طبقه قبل است. اگر  $f(x,y)$  تابع چگالی احتمال مشترک متغیرهای  $x$  و  $y$  باشد، در این صورت امید ریاضی شرطی  $y$  در  $x$  به صورت رابطه (۱۴) خواهد بود:

$$E(y|X) = \frac{\int_{-\infty}^{+\infty} yf(X,y)dy}{\int_{-\infty}^{+\infty} f(X,y)dy} \quad \text{رابطه (۱۴)}$$

معمولاً تابع چگالی احتمال براساس نمونه‌هایی از مشاهدات  $x$  و  $y$  تولید می‌شود. فرض کنید که  $\hat{f}(X,Y)$  یک برآوردکننده تابع چگالی احتمال براساس متغیرهای تصادفی  $x$  و  $y$  باشد. در این صورت خواهیم داشت:

$$\hat{f}(X,Y) = \frac{1}{(2\pi)^{\frac{p+1}{2\sigma^{p+1}}}} \frac{1}{n} \times$$

$$\prod_{i=1}^n \exp\left[-\frac{(X-X^i)^T(X-X^i)}{2\sigma^2}\right] \exp\left[-\frac{(Y-Y^i)^2}{2\sigma^2}\right] \quad \text{رابطه (۱۵)}$$

در رابطه (۱۵)،  $n$  تعداد مشاهدات نمونه،  $p$  بیانگر بعد بردار  $x$  و  $\sigma$  نشان‌دهنده پارامتر گسترش در مدل GRNN است. تابع اسکالر  $D_i^2$  و کرنل گوسی به صورت روابط (۱۶) و (۱۷) در نظر گرفته می‌شوند:

$$D_i^2 = (X - X^i)^T (X - X^i) \quad \text{رابطه (۱۶)}$$

$$K(X, X^i) = \exp\left(-\frac{D_i^2}{2\sigma^2}\right) \quad \text{رابطه (۱۷)}$$

با انجام محاسبات مورد نیاز، تابع برآوردگر  $\hat{Y}(X)$  به صورت رابطه (۱۸) حاصل خواهد شد:

$$\hat{Y}(X) = \frac{\sum_{i=1}^n Y^i K(X, X^i)}{\sum_{i=1}^n K(X, X^i)} \quad \text{رابطه (۱۸)}$$

یک مدل GRNN شامل چهار لایه است. لایه ورودی شامل  $p$  نرون است که براساس پارامترهای ورودی تعریف

رابطه خطی در فضای ویژگی معادل رابطه غیرخطی بین ورودی و خروجی در فضای اصلی است.

در این الگوریتم، هدف ایجاد تابعی است که به بهترین شکل داده‌ها را طبقه‌بندی کند و این تابع به کمک نمونه‌های آموزشی به دست می‌آید. طبقه‌بندی کننده‌های خطی زیادی وجود دارند که می‌توانند داده‌ها را از هم جدا کنند اما فقط یکی از این طبقه‌بندی کننده‌ها وجود دارد که برخلاف سایر روش‌ها که اغلب خطای مدل‌سازی را مینیمم می‌کنند، به جای آن ریسک عملیاتی تابع هدف را در نظر می‌گیرد و مقدار بهینه آن را محاسبه می‌کند. معمولاً تابع هزینه در مدل SVR را می‌توان به صورت رابطه (۱۲) بیان نمود:

$$L_\varepsilon(d, y(\mathbf{x})) = \begin{cases} |d - y(\mathbf{x})| - \varepsilon & \text{for } |d - y(\mathbf{x})| \geq \varepsilon \\ 0 & \text{for } |d - y(\mathbf{x})| < \varepsilon \end{cases} \quad \text{رابطه (۱۲)}$$

در رابطه (۱۲)،  $d$  بیانگر خروجی مطلوب،  $x$  نشان‌دهنده بردار مشاهدات ورودی،  $y(x)$  بیانگر خروجی حاصل از مدل SVR و  $\varepsilon$  نشان‌دهنده سطح خطای مجاز است. در حقیقت با استفاده از پارامتر  $\varepsilon$  بازه‌ای تعریف می‌شود که در داخل آن خطا قابل صرف نظر کردن است. خروجی مدل SVR را می‌توان با استفاده از رابطه (۱۳) محاسبه نمود (واپنیک، ۱۹۹۵):

$$f(\mathbf{x}_{SVR}) = y_{SVR} = \sum_{i=1}^M w_i x_i + b = \mathbf{w}^T \mathbf{x}_{SVR} + b \quad \text{رابطه (۱۳)}$$

در رابطه (۱۳)  $x_{SVR}$  بیانگر مشاهدات ورودی به مدل SVR،  $w$  نشان‌دهنده بردار وزن‌ها و  $b$  بیانگر بایاس است.

### ۳-۴- مدل شبکه عصبی رگرسیون عمومی

شبکه عصبی رگرسیون عمومی (GRNN) نوعی از شبکه‌های عصبی توابع پایه شعاعی (RBNN) است که در سال ۱۹۹۱ توسط اسپیچ ارائه شد (اسپیچ، ۱۹۹۱). این مدل برای رگرسیون، طبقه‌بندی و پیش‌بینی سیستم‌های دینامیک بسیار

#### ۴- نتایج عددی و بحث

در این بخش نتایج عددی و آنالیزهای انجام گرفته در مراحل آموزش و آزمون برای مدل‌های مختلف ارائه شده و نتایج حاصل، تفسیر می‌شوند.

#### ۴-۱- نتایج مرحله آموزش مدل‌های SVR, ANFIS, ANN و GRNN

هدف از مرحله آموزش در مدل‌های یادگیری ماشین، تعیین مدل بهینه و ساختار شبکه است تا براساس آن مدل دارای کمینه خطا شود. معمولاً دیتاهای مورد استفاده در مدل‌های یادگیری ماشین به سه دسته آموزش، اعتبارسنجی و تست تقسیم می‌شوند. هفتاد درصد دیتاها برای مرحله آموزش استفاده می‌شود. در این مرحله کمیت‌های ورودی و خروجی متناظر با آن به مدل معرفی شده و سپس، خروجی مدل، با خروجی مطلوب مقایسه می‌شود. براساس تابع هدف (رابطه (۲)) مقدار خطای مدل حاصل شده و سپس براساس آن، وزن‌ها و بایاس‌های مدل تصحیح می‌شوند. در این مقاله در مرحله آموزش دو پارامتر RMSE و ضریب همبستگی برای چهار مدل و در دو حالت محاسبه شده است. جدول (۱) مقادیر این دو پارامتر را در مرحله آموزش نشان می‌دهد.

براساس نتایج حاصل از جدول (۱) در هر دو حالت اول و دوم، خطای مدل ANFIS در مرحله آموزش از خطای سایر مدل‌ها کمتر است. این مدل ترکیبی از دو مدل ANN و FIS بوده و ویژگی‌های ساختاری هر دو را شامل می‌شود. در نتیجه آموزش در این مدل با دقت بالاتری نسبت به سایر مدل‌ها انجام می‌گیرد. همچنین نتایج جدول (۱) نشان می‌دهد که مدل SVR در هر دو حالت خطای زیادی داشته است. با توجه به اینکه در این مدل پارامترهای تابع کرنل براساس مشاهدات ورودی تعیین می‌شوند، پایین بودن تعداد مشاهدات ورودی در نتایج این مدل تأثیر زیادی خواهد داشت. به عبارت دیگر این مدل برای تعداد مشاهدات زیاد، مناسب است. مقادیر ضریب همبستگی همه مدل‌ها در هر

می‌شود. لایه الگو شامل  $n$  گره است که با نمونه‌های  $x^i$  فعال می‌شوند. در هر گره لایه الگو، از معادله (۱۶) برای محاسبه  $D_i^2$  استفاده شده و سپس مقدار حاصل در رابطه (۱۷) به کار گرفته می‌شوند. خروجی لایه الگو به لایه جمع‌بندی ارسال می‌شوند. داده‌های ورودی به همراه ضرایب مربوطه به تابع شعاعی ارائه می‌شوند. بعد از محاسبه خروجی این لایه، مقادیر خروجی با در نظر گرفتن ضرایب دیگری به لایه خطی داده می‌شوند. در نهایت خروجی نهایی از ترکیب خطی خروجی‌های لایه خطی به دست می‌آید. خوشه‌بندی داده‌های آموزش ورودی در لایه شعاعی انجام می‌شود. از این رو، تعداد نورون‌ها در لایه شعاعی برابر با تعداد مجموعه داده‌های مورد استفاده برای آموزش است. لایه خطی همیشه از نورون اضافی در مقایسه با لایه خروجی بهره می‌گیرد. این نورون اضافی، تابع چگالی احتمال را محاسبه می‌کند، در حالی که نورون‌های باقیمانده برای محاسبه خروجی استفاده می‌شود (خسروی و همکاران، ۱۳۹۹؛ غریتلار و پرساد، ۲۰۱۸).

#### ۳-۵- مدل ژئوئید محلی IRG2016

مدل ژئوئید محلی ایران تحت عنوان IRG2016، مدلی براساس توابع پایه شعاعی با تلفیق داده‌های زمینی و گرادئومتری ماهواره‌ای بوده که در سازمان نقشه‌برداری کشور و تحت رساله دکتری دانشگاه تهران انجام شده است (سعادت و همکاران، ۲۰۱۸). این مدل برای محدوده ایران از عرض جغرافیایی ۲۵ تا ۴۰ درجه و طول جغرافیایی ۴۴ تا ۶۳/۵ درجه محاسبه شده و در آن از ۲۱۵۲۵ داده شتاب‌ثقل پالایش شده توسط مرکز داده سازمان نقشه‌برداری کشور استفاده شده است. در این مدل دو روش برای تعیین ارتفاع ژئوئید نقاط در نظر گرفته شده است که یکی براساس درونیابی شبکه نقاط  $2.5^0 \times 2.5^0$  و دیگری براساس محاسبات مجزا با استفاده از مدل ژئوپتانسیلی جهانی و ضرایب توابع پایه شعاعی است. در روش محاسبات مجزا علاوه بر طول و عرض جغرافیایی، باید ارتفاع بیضوی نقطه مدنظر بر روی سطح زمین نیز به مدل معرفی شود.

جدول ۱: مقایسه مقادیر خطای RMSE (سانتی متر) و ضریب همبستگی مدل‌های ANN، ANFIS، SVR و GRNN در مرحله آموزش و دو حالت مختلف از ایستگاه‌ها

ضریب همبستگی				خطای RMSE (سانتی متر)				
GRNN	SVR	ANFIS	ANN	GRNN	SVR	ANFIS	ANN	
۰/۹۶	۰/۹۵	۰/۹۸	۰/۹۷	۵۴/۱۱	۵۹/۵۹	۸/۳۸	۳۰/۱۷	حالت اول (۲۲ ایستگاه آموزش)
۰/۹۸	۰/۹۵	۰/۹۷	۰/۹۷	۳۹/۶۷	۶۲/۱۱	۱۵/۹۴	۲۳/۷۷	حالت دوم (۱۹ ایستگاه آموزش)

نسبت به مدل‌های ANN، SVR و GRNN برخوردار است. بایستی اشاره شود که مدل محلی IRG2016 از مشاهدات کلیه ایستگاه‌های فلات ایران برای مدل‌سازی ژئوئید محلی استفاده می‌کند و بنابراین انتظار می‌رود که این مدل در منطقه مورد مطالعه دقت بالاتری نسبت به سایر مدل‌های مورد بررسی داشته باشد.

جدول ۲: مقادیر خطای نسبی (درصد) مدل‌های ANN، ANFIS، SVR و GRNN در ایستگاه‌های آزمون در نظر گرفته شده برای حالت اول

AHAR	SAGZ	BZGN	TABZ	
۰/۰۴	۲/۸۶	۱/۶۸	۰/۲۵	IRG2016
۲/۵۶	۳/۶۴	۰/۸۶	۱/۱۰	ANN
۰/۹۶	۱/۹۴	۰/۸۹	۰/۱۰	ANFIS
۲/۲۹	۴/۶۶	۲/۳۶	۰/۷۱	SVR
۲/۰۶	۵/۹۲	۱/۸۳	۰/۷۴	GRNN

بر اساس نگاره (۱) ایستگاه آزمون SAGZ از سایر ایستگاه‌های آموزش فاصله زیادی دارد. مقایسه خطای نسبی مدل‌ها در این ایستگاه با سایر ایستگاه‌های آزمون، نشان‌دهنده خطای بیشتر همه مدل‌ها در این ایستگاه است. اما، ایستگاه آزمون TABZ با ایستگاه‌های آموزش زیادی احاطه شده است. خطای نسبی مدل‌ها در این ایستگاه نسبت به سایر ایستگاه‌های آزمون خیلی پایین‌تر است. به عبارت دیگر، خطای مدل‌های یادگیری ماشین به دیتاهای

دو حالت، بالای ۰/۹۵ است. به عبارت دیگر، همه مدل‌ها تغییرات موجود در ارتفاع ژئوئید را به درستی نشان می‌دهند. ساختار بهینه انتخاب شده برای مدل ANN به صورت ۱-۳-۲ بوده و برای مدل ANFIS به صورت ۱-۳-۲ است. در این ساختارها، عدد ۱ نمایش‌دهنده خروجی مدل یعنی ارتفاع ژئوئید بوده و عدد ۲ بیانگر پارامترهای ورودی یعنی طول و عرض جغرافیایی است. در مدل ANN عدد ۴ بیانگر تعداد نوروهای لایه پنهان و در مدل ANFIS نشان‌دهنده تعداد قوانین اگر-آنگاه فازی است.

#### ۴-۲- نتایج مرحله آزمون مدل‌ها

پس از مرحله آموزش مدل‌های یادگیری ماشین و انتخاب ساختار بهینه برای آن‌ها، در دو حالت مختلف (۴ و ۷ ایستگاه) مرحله آزمون انجام می‌گیرد. در این مرحله مقدار ارتفاع ژئوئید در ایستگاه‌های آزمون برآورد شده و با مقدار حاصل از تفاضل ارتفاع اورتومتريک و ارتفاع نرمال به عنوان مبنا، مقایسه می‌شود. دو شاخص آماری خطای نسبی برحسب درصد و خطای RMSE برحسب سانتی‌متر برای تمامی مدل‌ها محاسبه شده و در جداول (۲) و (۳) ارائه شده است.

نتایج جدول (۲) و مقایسه مقادیر خطای نسبی همه مدل‌ها در ایستگاه‌های آزمون، نشان می‌دهد که مدل ANFIS نسبت به سایر مدل‌ها از دقت بالاتری برخوردار بوده است. پس از مدل ANFIS، مدل IRG2016 از دقت بالاتری

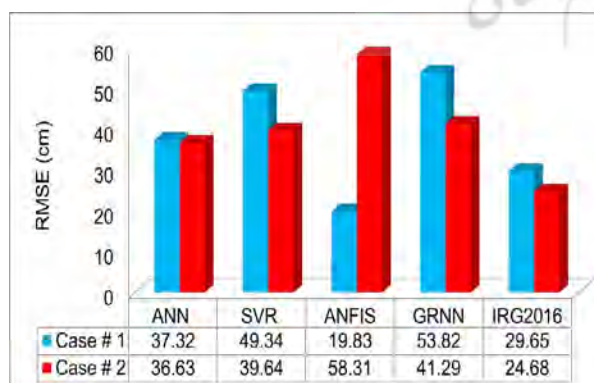


جدول ۳: مقادیر خطای نسبی (درصد) مدل‌های ANN، ANFIS، SVR و GRNN و IRG2016 در ایستگاه‌های آزمون در نظر گرفته شده برای حالت دوم

BNAB	ARDH	VLDN	AHAR	SAGZ	BZGN	TABZ	
۰/۹۸	۰/۶۸	۰/۹۷	۰/۰۴	۲/۸۶	۱/۶۸	۰/۲۵	<b>IRG2016</b>
۱/۳۴	۳/۵۹	۳/۲۰	۰/۶۴	۰/۵۰	۱/۹۷	۰/۳۲	<b>ANN</b>
۰/۹۰	۲/۷۲	۱/۲۰	۶/۸۳	۵/۸۸	۱/۵۶	۰/۵۵	<b>ANFIS</b>
۲/۲۲	۲/۶۸	۰/۷۸	۲/۲۵	۴/۲۷	۱/۹۵	۰/۷۵	<b>SVR</b>
۱/۳۱	۳/۹۱	۰/۴۲	۱/۸۸	۵/۴۶	۰/۰۵	۰/۷۰	<b>GRNN</b>

در حالت اول، مدل ANFIS نسبت به سایر مدل‌های بررسی شده خطای نسبی کمتری داشته ولی در حالت دوم و با کاهش تعداد ایستگاه‌های آموزش، مدل محلی ژئوئید IRG2016 دقت بالایی را ارائه کرده است. البته در مقایسه مابین مدل‌های یادگیری ماشین در حالت دوم، مدل ANN خطای نسبی کمتری داشته است. در حالت کلی و با توجه به اینکه مدل ANFIS ترکیبی از مدل‌های ANN و FIS است، می‌توان نتیجه گرفت که این مدل قابلیت بسیار بالایی در برآورد ارتفاع ژئوئید محلی با استفاده از مشاهدات طول و عرض جغرافیایی در منطقه مورد مطالعه دارد.

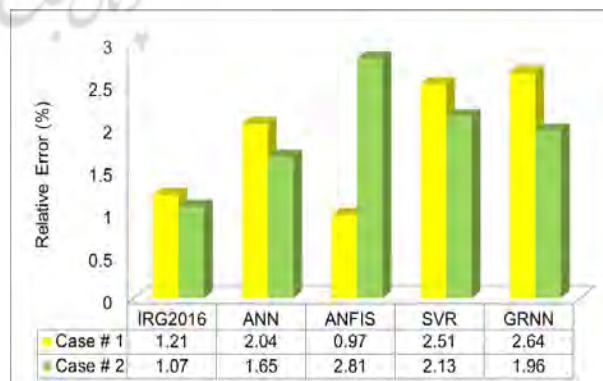
نگاره (۷) مقادیر خطای RMSE (برحسب سانتی‌متر) مدل‌های ANN، ANFIS، SVR، GRNN و IRG2016 را در مرحله آزمون نشان می‌دهد.



نگاره ۷: مقایسه مقادیر RMSE (بر حسب سانتی‌متر) مدل ANN، ANFIS، SVR، GRNN و IRG2016 در مرحله آزمون

ورودی و چگونگی پخش و توزیع ایستگاه‌های آموزش بسیار وابسته هستند. با این وجود مدل ANFIS در ایستگاه SAGZ کمترین خطای نسبی را داشته است. جدول (۳) نتایج خطای نسبی محاسبه شده همه مدل‌های مورد بررسی در این مقاله را برای حالت دوم یعنی ۷ ایستگاه آزمون نشان می‌دهد. در حالت دوم تعداد ایستگاه‌های آموزش نسبت به حالت اول کمتر شده است.

در حالت دوم نیز، بیشترین مقدار خطای نسبی در ایستگاه آزمون SAGZ که فاصله زیادی با ایستگاه‌های آموزش دارد و کمترین خطا در ایستگاه آزمون TABZ تولید شده است. نتایج جدول (۳) نشان می‌دهد که با کاهش تعداد ایستگاه‌های آموزش، مدل ANN دقت بالاتری را ارائه داده است. نگاره (۶) میانگین خطای نسبی همه مدل‌های مورد بررسی را در دو حالت نشان داده است.



نگاره ۶: میانگین خطای نسبی مدل‌های ANN، IRG2016، ANFIS، SVR و GRNN در دو حالت و در ایستگاه‌های آزمون



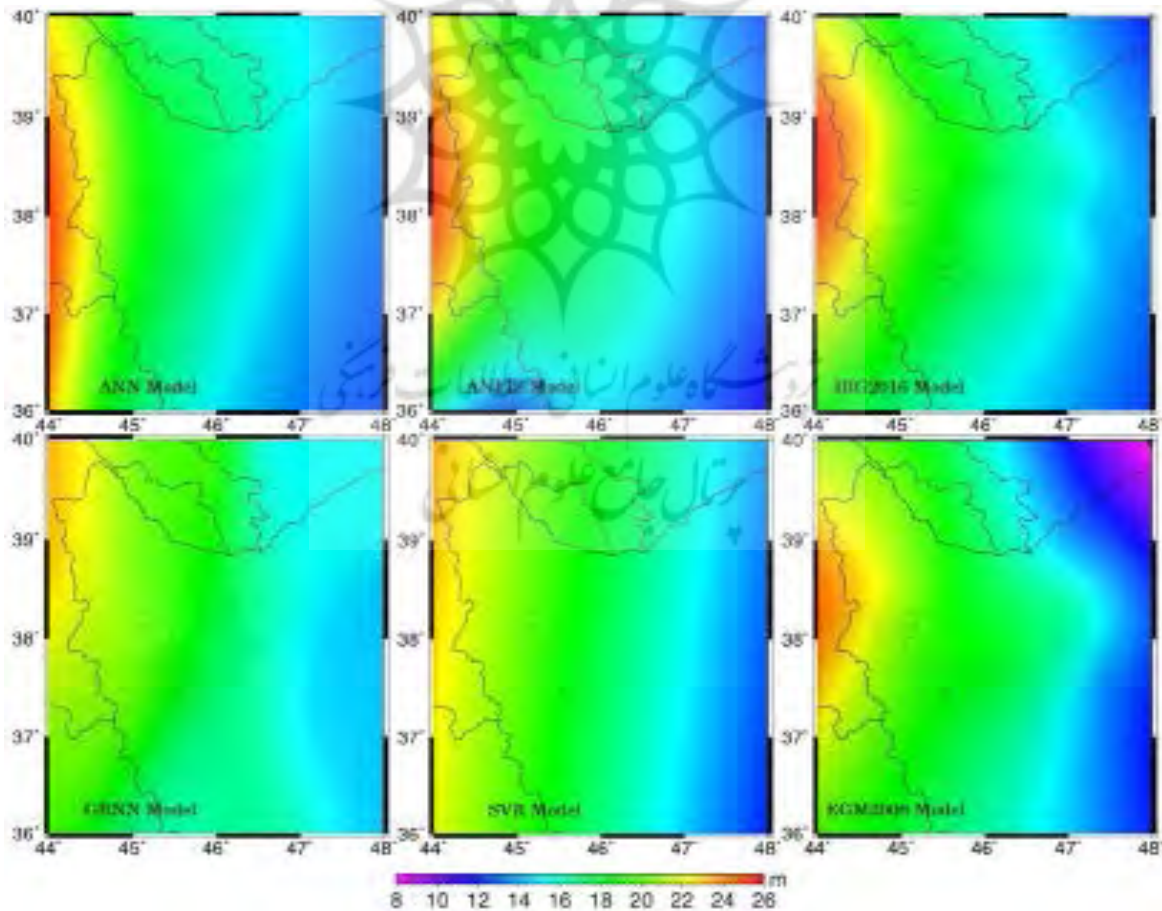
در حالت اول و مطابق با نگاره (۷) خطای مدل ANFIS،

۱۹/۸۳ سانتی متر شده که در مقایسه با سایر مدل‌ها، دقت بالاتری را دارد. اما در حالت دوم مدل IRG2016 دارای خطای ۲۴/۶۸ سانتی متر است. مقایسه نتایج در دو حالت نشان می‌دهد که با تعداد ایستگاه‌های آموزش بیشتر، مدل ANFIS دقت بالاتری داشته و با تعداد ایستگاه‌های آموزش کمتر، مدل ANN نسبت به سایر مدل‌های یادگیری ماشین بررسی شده در این مقاله، خطای کمتری دارد. البته بایستی مجدداً اشاره نمود که مدل محلی IRG2016 در حالت دوم دقت بالاتری داشته است اما، این مدل از اندازه‌گیری‌های کلیه ایستگاه‌های فلات ایران برای مدل‌سازی استفاده می‌کند. مقایسه‌های انجام گرفته با این مدل بیشتر برای نشان دادن عیار دقت مدل‌های یادگیری ماشین در برآورد ارتفاع

ژئوئید محلی در منطقه مورد مطالعه است.

#### ۴-۳- برآورد ارتفاع ژئوئید در منطقه مورد مطالعه

پس از مرحله آموزش و آزمون مدل‌های ANN، ANFIS، SVR، GRNN و IRG2016 و مشخص شدن دقت تمامی مدل‌ها، حال می‌توان مقدار ارتفاع ژئوئید را در محدوده شبکه مورد مطالعه برآورد نمود. در مدل‌های یادگیری ماشین مورد بررسی در این مقاله، با ارائه طول و عرض هر نقطه دلخواه ژئودیتیکی داخل شبکه مورد نظر، ارتفاع ژئوئید برآورد می‌شود. نگاره (۸) مقادیر ارتفاع ژئوئید برآورد شده در محدوده شبکه مورد مطالعه را با استفاده از مدل‌های مورد نظر نشان می‌دهد. در این نگاره و به منظور مقایسه چگونگی تغییرات ارتفاع ژئوئید، خروجی



نگاره ۸: چگونگی تغییرات ارتفاع ژئوئید برآورد شده توسط مدل‌های ANN، ANFIS، IRG2016 (ردیف بالا)، EGM2008

و SVR (ردیف پایین) در محدوده شبکه مورد مطالعه در این مقاله

همکاران برای منطقه ایران ارائه شده و همچنین مدل ژئوئید جهانی EGM2008 مقایسه شد.

ارزیابی‌های آماری انجام گرفته نشان داد که در حالت ۲۲ ایستگاه آموزش و ۴ ایستگاه آزمون، RMSE مدل‌های ANN، ANFIS، SVR، GRNN و IRG2016 در مرحله آزمون به ترتیب برابر با ۳۷/۳۲، ۱۹/۸۳، ۴۹/۳۴، ۵۳/۸۲ و ۲۹/۶۵ سانتی‌متر می‌شوند. اما در حالت ۱۹ ایستگاه آموزش و ۷ ایستگاه آزمون، مقادیر خطای مدل‌ها به ترتیب برابر با ۳۶/۶۳، ۵۸/۳۱، ۳۹/۶۴، ۴۱/۲۹ و ۲۴/۶۸ سانتی‌متر شده است. به عبارت دیگر با کاهش تعداد ایستگاه‌های آموزش، خطای مدل‌های یادگیری ماشین افزایش می‌یابد. نتایج حاصل از این مقاله نشان داد که با وجود تعداد ایستگاه‌های آموزش کافی، مدل‌های ANFIS و ANN از دقت بسیار بالایی در برآورد ژئوئید محلی برخوردار هستند و می‌توانند جایگزین مدل‌های جهانی همانند EGM2008 شوند. همچنین از این مدل‌ها می‌توان در کارهای عمرانی و نقشه‌برداری نیز بهره برد. برای ادامه این تحقیق می‌توان از تعداد نقاط بیشتر و متراکم‌تر به منظور بررسی دقت مدل‌های یادگیری ماشین استفاده کرده و همچنین ورودی مدل‌های یادگیری ماشین را براساس پارامترهای هندسی و فیزیکی در نظر گرفته و دقت را ارزیابی نمود.

#### منابع و مأخذ

۱- خسروی، م.، حسینی، ا.، قاسمی، ج.، (۱۳۹۹). ارزیابی عملکرد منابع انسانی با رویکرد شبکه عصبی رگرسیون عمومی، مورد مطالعه: اعضای هیئت علمی. دو فصلنامه مطالعات برنامه‌ریزی آموزشی، دوره نهم شماره هفدهم، بهار و تابستان ۱۳۹۹.

۲- خضرائی، س.م.، نفیسی، و.، منجمی، س.ا.، عسگری، ج.، امیری سیمکویی، ع.، (۱۳۹۳). مدل‌سازی ژئوئید محلی دقیق با استفاده از داده‌های GPS، ترازیبی و تکنیک‌های هوش مصنوعی: مطالعه موردی شاهین‌شهر اصفهان. نشریه علمی پژوهشی علوم و فنون نقشه‌برداری ۴ (۳): ۲۳۸-۲۲۵.

مدل ژئوئید جهانی EGM2008 نیز نمایش داده شده است. مطابق با نتایج نگاره (۸) و خروجی مدل‌های IRG2016 و EGM2008، در منطقه مورد مطالعه ارتفاع ژئوئید در عرض جغرافیایی ۳۷ الی ۴۰ و طول جغرافیایی ۴۴ الی ۴۵ درجه دارای بیشترین تغییرات است. با دور شدن از محدوده مورد اشاره، ارتفاع ژئوئید کمتر می‌شود. مقایسه نتایج مدل‌های ANFIS و ANN با خروجی مدل‌های IRG2016 و EGM2008 نشان می‌دهد که این دو مدل یادگیری ماشین، تغییرات ارتفاع ژئوئید را از لحاظ کمی و کیفی با دقت بالایی ارائه می‌دهند. اما خروجی دو مدل SVR و GRNN از لحاظ مقدار عددی تا حدودی با مدل‌های IRG2016 و EGM2008 متفاوت است. بنابراین نتیجه حاصل از نگاره (۸) نشان می‌دهد که مدل‌های ANFIS و ANN دارای دقت بالایی در برآورد ارتفاع ژئوئید محلی هستند و می‌توانند جایگزین مدل‌های جهانی و محلی شوند.

#### ۵- نتیجه‌گیری

در این مقاله هدف برآورد ارتفاع ژئوئید محلی با استفاده از مدل‌های یادگیری ماشین بود. برای انجام این کار از مدل‌های شبکه عصبی مصنوعی (ANN)، مدل استنتاج عصبی-فازی سازگار (ANFIS)، رگرسیون بردار پشتیبان (SVR) و مدل شبکه عصبی رگرسیون عمومی (GRNN) استفاده شد. مختصات ژئودتیکی ۲۶ ایستگاه در شبکه شمال غرب ایران به همراه ارتفاع اورتومتريک ( $H_0$ ) و ارتفاع نرمال (h) آن‌ها از سازمان نقشه‌برداری کشور تهیه شد. در تمامی ایستگاه‌ها، تفاضل ارتفاع اورتومتريک و ارتفاع نرمال به عنوان ارتفاع ژئوئید (N) در آن نقطه در نظر گرفته شد. بنابراین، طول و عرض ژئودتیکی ایستگاه‌های GPS به عنوان ورودی مدل‌های یادگیری ماشین بوده و ارتفاع ژئوئید متناظر با آن، به عنوان خروجی در نظر گرفته شد. به منظور تست نتایج مدل‌های یادگیری ماشین، دو حالت ۴ و ۷ ایستگاه آزمون در نظر گرفته شد. همچنین خروجی مدل‌ها با مدل ژئوئید محلی IRG2016 که توسط سعادت و

- E. (eds) Encyclopedia of Geodesy. Springer, Cham. [https://doi.org/10.1007/978-3-319-02370-0\\_44-1](https://doi.org/10.1007/978-3-319-02370-0_44-1).
- 14- Kiamehr, R., (2006). Precise Gravimetric Geoid Model for Iran Based on GRACE and SRTM Data and the Least-Squares Modification of Stokes' Formula: with Some Geodynamic Interpretations. PhD Thesis. Royal Institute of Technology, Stockholm, Sweden.
- 15- Mars, P., J.R. Chen, R. Nambiar (1996). Learning Algorithms: Theory and Applications in Signal Processing, Control and Communications, CRC Press, Boca Raton, Florida.
- 16- Nahavandchi, H., Soltanpour, A., (2005). Improved determination of heights using a conversion surface by combining gravimetric quasi/geoid and GPS-levelling height differences. Stud. Geophys. Geod., 50, 165-180.
- 17- Najafi M., (2004). Determination of Precise Geoid for Iran Based on Stokes-Helmert Scheme. Report 2003. TOTAK Project, National Cartographic Center of Iran (NCC), Tehran, Iran.
- 18- Saadat, A., Safari, A., Needell, D., (2018). IRG2016: RBF-based regional geoid model of Iran, Studia Geophysica et Geodaetica, 62, DOI: 10.1007/s11200-016-0679-x.
- 19- Sabri, L.M., Sunantyo, L, Widjajanti, N., (2019). Geoid determination with Hotine's integral based on terrestrial gravity data in Semarang city. IOP Conf. Series: Journal of Physics: Conf. Series 1127 (2019) 012047 IOP Publishing doi:10.1088/1742-6596/1127/1/012047.
- 20- Safari A., Ardalan A.A., Grafarend E.W., (2005). A new ellipsoidal gravimetric, satellite altimetry and astronomic boundary value problem, a case study: The geoid of Iran. J. Geodyn., 39, 545-568.
- 21- Sedighi M., Najafi-Alamdari M. and Tabatabaie S.H., (2008). Gravity field implied density modeling, for precise determination of the geoid. J. Appl. Sci., 8, 3371-3379.
- 22- Specht, D. F., (1991). A general regression neural network. IEEE Transactions on Neural Networks, vol. 2, no. 6, pp. 568-576, Nov. 1991, doi: 10.1109/72.97934.
- 23- Vaníček, P., Krakiwsky, E., (1986). Geodesy, the Concepts. North Holland, 1986, ISBN: 0444877770, 9780444877772.
- 3- Ardalan A.A., (1999). High Resolution Regional Geoid Computation in the World Geodetic Datum 2000 Based upon Collocation of Linearized Observational Functionals of the Type Gravity Potential and Gravity Intensity. PhD Thesis. Stuttgart University, Stuttgart, Germany, 250 pp.
- 4- Ardalan, A.A., Hatam, C.Y., Sharifi, M.A., Safari, A., Gazavi, K., Motagh M., (2002). Determination of Precise geoid for Iran based on Potential approach. Technical report, National Cartographic Center of Iran (NCC), Tehran, Iran.
- 5- Foroughi, I., Goli, m., Pagiatakis, S., Ferguson, S., Novák, P., (2023). Data requirements for the determination of a sub-centimetre geoid. Earth-Science Reviews, (239), April 2023, 104326.
- 6- Ghritlahre, H., Prasad, R., 2018, Investigation of thermal performance of unidirectional flow porous bed solar air heater using MLP, GRNN, and RBF models of ANN technique. Thermal Science and Engineering Progress. doi: <https://doi.org/10.1016/j.tsep.2018.04.006>.
- 7- Hamesh, M., Zomorrodian, H., (1992). Iranian gravimetric geoid determination, second step. NCC J. Surv., 6, 17-24, 52-63.
- 8- Hatam, C.Y., (2010). Etablissement des nouveaux reseaux multi-observations geodesiques et gravimetriques et determination du geoide en Iran. PhD Thesis. Geophysics, University Montpellier 2, Montpellier, France (in French).
- 9- Haykin, S. (1994): Neural Networks, a comprehensive Foundation, Macmillan College Publishing Company, New York.
- 10- Hofmann-Wellenhof, B., Moritz, H., (2006). Physical Geodesy. Springer Vienna, <https://doi.org/10.1007/978-3-211-33545-1>.
- 11- Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. IEEE transactions on systems, man, and cybernetics, Vol. 23, No.3, pp.665-685.
- 12- Kavzoglu, T., Saka, M. (2005). Modeling local GPS/leveling geoid undulations using artificial neural networks. J Geodesy 78, 520-527, doi.org/10.1007/s00190-004-0420-3.
- 13- Kenyeres, A., (2016). GPS/Leveling. In: Grafarend,

24- Vapnik, V. (1995). Support-vector Networks. Machine Learning, 20, 273-297.

25- Weber, G., Zomorrodian, H., (1988). Regional geopotential model improvement for the Iranian geoid determination. Bull. Géod., 62, 125-141.

26- Yeganeh, B., Motlagh, M. S. P., Rashidi, Y., Kamalan, H. (2012). Prediction of CO concentrations based on a hybrid Partial Least Square and Support Vector Machine model. Atmospheric Environment, 55, 357-365.

27- Zaletnyik, P., Völgyesi, L., Paláncz, B., (2007). Modeling local GPS/leveling geoid undulations using Support Vector Machines. Civil Engineering 52/1 (2008) 39-43, doi: 10.3311/pp.ci.2008-1.06.



---

## COPYRIGHTS

©2024 by the authors. Published by National Geographical Organization. This article is an open access article distributed under the terms and conditions of the Creative Commons [Attribution-NoDerivs 3.0 Unported \(CC BY-ND 3.0\)](https://creativecommons.org/licenses/by-nd/3.0/)

---





پروہشگاہ علوم انسانی و مطالعات فرہنگی  
پرتال جامع علوم انسانی