داده کاوی بر روی داده های هواشناسی و

پیاده سازی نرم افزار آن برای شبیه سازی پرواز.

نگارنده

ابوالفضل قهرمانی

دانشجوی کارشناسی ارشد دانشگاه صنعتی امیرکبیر تهران

|  |  |
| --- | --- |
| فهرست مطالب | صفحه |

[فصل اول 4](#_Toc534414292)

[مروری بر آثار ، تحقیق و پژوهش جهانی در زمینه پیش‌بینی آب‌وهوا 5](#_Toc534414293)

[اقلیم و آب و هوا، شناخت و ارتباط بین آن دو 5](#_Toc534414294)

[دانش اتصال دهنده آب‌وهوا و اقلیم جغرافیایی به یکدیگر 6](#_Toc534414295)

[چگونه آب و هوا و اقلیم جغرافیایی بهم مرتبط می شوند؟ 6](#_Toc534414296)

[Jet streams 7](#_Toc534414297)

[اتحادیه NOAA 11](#_Toc534414298)

[لابراتوار NOAA Earth System Research 12](#_Toc534414299)

[بازتحلیلی هواشناختی(طراحی سامانه درک و فهم داده های اقلیم) 13](#_Toc534414300)

[مدل NCEP/NCAR Reanalysis 13](#_Toc534414301)

[Accessing the data 13](#_Toc534414302)

[مدل CISL Research Data Archive 14](#_Toc534414303)

[مدل ECMWF re-analysis 14](#_Toc534414304)

[فصل اول تعریف مسئله 15](#_Toc534414305)

[تعریف مسئله 16](#_Toc534414306)

[فصل دوم دریافت و ذخیره داده ها از وبگاه weather.uwyo.edu 17](#_Toc534414307)

[دریافت و ذخیره داده ها 18](#_Toc534414308)

[تعریف پارامترهای فیزیکی 21](#_Toc534414309)

[فصل سوم تحلیل داده ها 24](#_Toc534414310)

[تحلیل داده ها 25](#_Toc534414311)

[ارزیابی از داده ها 26](#_Toc534414312)

[میان یابی درارتفاع موردنظر در توابع اتمسفر 28](#_Toc534414313)

[شبکه های عصبی بر روی داده ها 29](#_Toc534414314)

[تشریح شبکه عصبی و الگوریتم پیاده سازی آن 31](#_Toc534414315)

[تابع آموزش شبکه 33](#_Toc534414316)

[الگوریتم لونبرگ-مارکارد 33](#_Toc534414317)

[فصل چهارم شبیه ساز پرواز 36](#_Toc534414318)

[شبیه ساز پروازی 37](#_Toc534414319)

[زمان بندی تقریبی مراحل ساخت نرم افزار 37](#_Toc534414320)

# فصل اول

**تاریخچه، پژوهش ها، روش ها و نوع مظالعه بر روی اتمسفر کره زمین**

# **مروری بر آثار ، تحقیق و پژوهش جهانی در زمینه پیش**‌بینی آب‌وهوا

از گذشته، با توجه به نیاز بشر و همچنین پیشرفت روزافزون علم در دهه های گذشته، مطالعه روی پیش بینی جو زمین اهمیت فراوانی پیدا کرده است. فهم پدیده های حاکم بر جو کره زمین، این توانایی را برای بشر به وجود می آورد که حتی قادر باشد برای زمان آینده، آب و هوای منطقه ای خاص را با دقت قابل قبول و کاربردی ای، پیش‌بینی کند. سعی شده است این مطالعات در ادامه بصورت کامل آورده شود.

# اقلیم و آب و هوا، شناخت و ارتباط بین آن دو

آب و هوا چیزی است که انسان هر روز در بیرون ، آن را تجربه می کند ، این در حالی است که اقلیم، توصیفی است از آنچه که در یک مدت متناوب مثلا هفته ها و یا در یک ماه و یا در یک سال، انتظار می‌رود رخ دهد.

شرایط آب و هوا برای یک منطقه خاص، می تواند تحت تاثیر عواملی چون تنوع اقلیمی آن منطقه از قبیل اقیانوس های گرمسیری،جو زمین و یا حتی گردابه های قطبی قرار بگیرد.این شرایط حتی می تواند تحت تاثیر مقدار رطوبت موجود در خاک، دمای حاکم بر سطح دریای منطقه، برف باریده موجود بر روی زمین و ...باشد.

## دانش اتصال دهنده آب‌وهوا و اقلیم جغرافیایی[[1]](#footnote-1) به یکدیگر

این دانش، برای بدست آوردن توانایی پیش بینی ارتباطات متقابل پدیده های اقلیم و آب و هوا و تبدیل این دانش به یک محصول عملیاتی پیش‌گو، ایچاد شده است.این مظالعات بیشتر در آژانس NAOO[[2]](#footnote-2) انجام می گردد.

برای این دانش و پژوهش دو جنبه بسیار مهم به وجود می آید، اول آن که تغییرات آرام اقلیم موجود (از قبیل جابجایی هفته به هفته jet stream ، تغییر در قدرت گردابه های قطبی و رفتار "رودخانه های اتمسفر" )که شدت و فرکانس سیستم های آب و هوایی را تخت تاثیر قرار می دهد، می بایست مورد مطالعه قرار بگیرد. دوم این که این دانش بدست آمده می بایست در سیستم های پیش بینی آب و هوا (NWS[[3]](#footnote-3)) و اقلیم مورد استفاده قرار بگیرد تا هشدارهایی برای رویدادهای شدید آب و هوایی برای عموم ارائه گردد.

## چگونه آب و هوا و اقلیم جغرافیایی بهم مرتبط می شوند؟

راه های زیادی برای مرتبط کردن این دو وجود دارد. بطور مثال تعداد تکرار و شدت طوفان های هاریکین تحت تاثیر تغییرات در دمای سطح دریا، الگوهای باد در لایه های بالاتر درتمام اقیانوس اطلس که بصورت ماهانه، فصلی، دهه به دهه،و سایر زمان بندی های ممکن، اتفاق می افتد.

تغییرات مکان jet stream ها در عرض جغرافیایی میانی می تواند منجر به ایجاد توفان ها شود.توفان های عرض میانی، به وسیله این جت استریم ها هدایت و کنترل می شوند، و انرژی مربوط به خود را از اختلاف دمایی هم بدست می آورند.

به طور نسبی تغییرات آرام در خصوصیات پدیده های جوی و الگو های آن بسیار قابل پیش بینی تر از سایر سیستم های هواشناسی می باشد.

ESRL[[4]](#footnote-4) توسعه یک نمونه از این کار را پیش رو گرفته است، که آن را رودخانه جوی [[5]](#footnote-5) نامیده است. رودخانه های جوی، ناحیه های باریکی از جو هستند که قسمتی از گردابه های فرا تروپی را تشکیل می دهند . این رودخانه ها مقدار عظیمی از بخار آب را در بر دارند ،که با خود آن را حمل می کند.تحقیقات نشان داده است که بیش از 90 % بخار آب های قطبی در ناحیه عرض جغرافیایی میانی قرار گرفته است، به علاوه ، وقتی با کوه های ساحلی بر خورد داشته باشد، منجر به باران های شدید و آوار می شود.به طور خلاصه این پدیده (رودخانه های جوی) نه تنها مقدار آب جهانی را تعیین می کند بلکه در شدت بارش ها نقش اساسی دارد . بدین ترتیب این موجودیت واسط اصلی اتصال اقلیم و آب و هوا را بر عهده می گیرد.

# Jet streams

جت استریم ها برای اولین بار، در خلال جنگ جهانی دوم کشف شدند(البته کشف هایی نیز از قبل رخ داده بود که مهم تلقی نشده بود). وقتی خلبان ها بطور دائم مسیر انگلستان تا آمریکا را پرواز می کردند، اذعان داشتند که پرواز به سمت انگلستان سریع تر انجام می شود، حدود 100 مایل بر ساعت سرعت این باد برای آنها ثبت می شد. این بادها در نوارهای باریکی در حال وزیدن هستند که آن ها را جت استریم می نامند.

جت استریم ها رودخانه های باریک هوا با سرعت بسیار زیاد هستند. علت تشکیل این جت ها به دو عامل جدا از هم مرتبط می شود. اولی وجود اختلاف دمای ناحیه استوایی- گرمسیری که با تابش خورشید بشدت گرم می شود، با ناحیه قطبی-سرد عرض های بالاتر زمین می باشد، که منجر به گرادیان فشار شدید می شود که باعث می شود هوا را به حرکت وا دارد این انتقال حرارت مدلی به اسم سلول های چرخش سه گانه (Polar, Ferrel, Hadley) ایجاد می کند، عامل دوم وجود شتاب یا نیرو کرولیوس زمین می باشد که ناشی از چرخش زمین به دور خود است. این عامل باعث می شود جریان ایجاد شده از توده های سرد و گرم جهت مستقیمی از گرم به سرد به خود نگیرد و با اثر کرولیوسی در میان مرزهای این توده هوا جریان پیدا کند.

هر نوار از جت استریم ها می تواند به کار خود پایان دهد و یا دوباره ایجاد شود، می تواند به چندین جت دیگر شکافته شود و یا اینکه باهم ادغام شوند.و در جهات مختلفی حتی خلاف جریان باقی مانده می توانند جریان پیدا کنند.

جت های قطبی قوی ترین جت استریم ها شناخته می شوند، که در ارتفاع 9 تا 12 کیلومتری از سطح دریا قرار گرفته اند،دسته دیگر جت های فروگرمسیری که ضعیف تر هستند و در ارتفاع 10 تا 16 کیلومتری سطح دریاها پیدا می شوند. جت قطبی های معمولا در عرضهای جغرافایی میانی 30 تا 60 درجه درجه(نزدیکه به 60 درجه بیشتر) قرار گرفته است این در حالی است که جت فروگرمسیری در عرض های جغرافیایی نزدیک 30 درجه حاکم هستند .

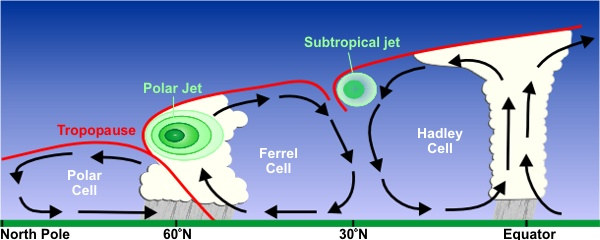
همان طور که در شکل 1 نشان داده شده است، مسیر این جت ها معمولا به صورت مورب و موج گونه می باشد.



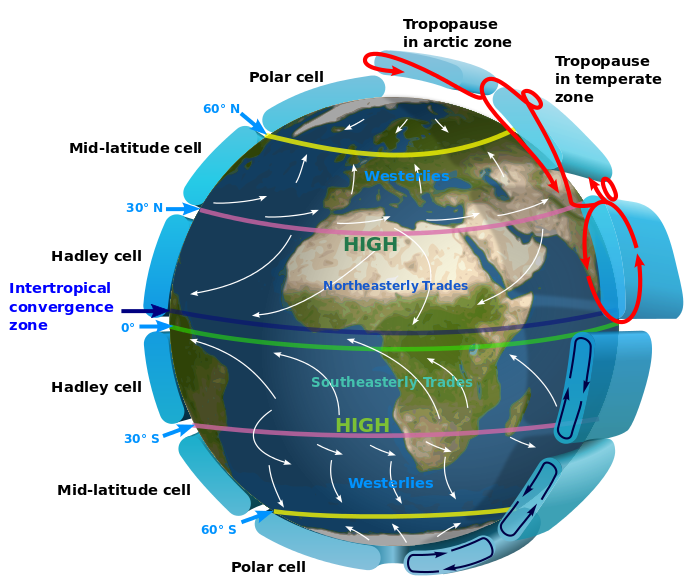
شکل 1نمایش جت استریم های دوگانه قطبی و فروگرمسیری

عامل اختلاف دمایی که ذکر شد، اتمسفر را به سه ناحیه اقلیمی تبدیل می کند، این ناحیه ها به نام های

(Polar, Ferrel, Hadley) شناخته شده اند، جت های استریم در مرز این سلول ها قرار دارند.(شکل های 2 و 3)

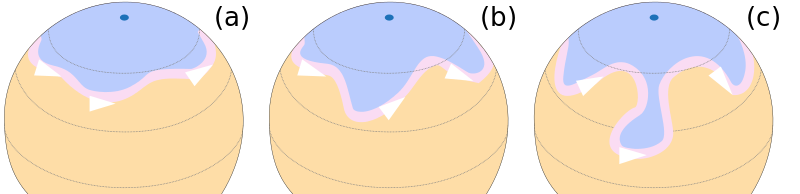


شکل 2برش عرضی از اتمسفر حاکم بر زمین، محوده های دو نوع جت استریم نمایش داده شده است.



شکل 3 جریان های چرخشی حاکم بر اتمسفر زمین که هر نیم کره را به سه لایه مختلف تبدیل می کند

جت استریم ها به صورت خطوط منحنی دار و به شکل موج های تاب دار در می آیند، ضخیم ترین این منحنی ها به موج های راسبی معروف اند، این جت ها امکان دارد بصورت توده ای از جت اصلی جدا شده و به عرض های پایین تر جبه هوای سرد را منتقل کنند(شکل 4).



شکل 4حرکت موج راسبی، در مرحله c قسمتی از جت قطبی کنده شده است و با خود هوای سرد را به عرض های پایین تر انتقال می دهد، مثالی از تاثیر جت استریم بر آب و هوا مناطق و توانایی در پیش بینی آن

با شناخت جت استریم ها، شناخت سیستم حاکم آب و هوا پیشرفت فراوانی کرده است، به طوری که قدرت پیش بینی پیامد ها و توفان های سهم گین و بارش های سیل آسا تا حدی بدست آمده است.

## اتحادیه NOAA

اتحادیه "**مدیریت جوی و اقیانوسی ملی** " که به اختصار **NOAA**[[6]](#footnote-6) نامیده می شود،این اتحادیه درباره رخداد پیامد های خطرناک آب و هوا هشدار می دهد و راهنمایی هایی برای استفاده از اقیانوس ها ارائه می دهد. همچنین نظارت بر محیط اتمسفر از کارهای این اتحادیه می باشد.

در سال 1970 این اتحادیه به صورت رسمی تشکیل شد، و تا سال 2017 حدود 11000 کارمند غیرنظامی را مشغول به کار کرده است.

این اتحادیه از سال 2017 به دست معاون بخش اتمسفر و اقیانوس وزارت بازرگانی مدیریت می شود.

NOAA نقش های خاصی در جامعه ایفا می کند که مزایای آن فراتر از اقتصاد ایالات متحده در سطح جامعه جهانی رسیده است. از جمله این نقش ها:

* یک تامین کننده اطلاعات محصولات زیست محیطی. NOAA به مشتریان و شرکای خود اطلاعات مربوط به وضعیت اقیانوس ها و جو را ارائه می دهد. روشن است این اطلاعات از طریق تولید هشدارها و پیش بینی های آب و هوایی از طریق سرویس هواشناسی ملی نیز به دست می آمد، اما محصولات NOAA به اطلاعات آب و هوا، اکوسیستم ها و تجارت نیز گسترش می یابد.
* ارائه دهنده خدمات مراقبت از محیط زیست
* نقش یک رهبر در تحقیقات علمی کاربردی.
* NOAA در نظر گرفته شده است که منبع اطلاعات دقیق و علمی در چهار حوزه خاص اهمیت ملی و جهانی که در بالا ذکر شده است: اکوسیستم، آب و هوا،اقلیم شناسی ، تجارت و حمل و نقل

پنج "فعالیت اساسی این اتحادیه " عبارتند از:

* نظارت و مشاهده سیستم های زمینی با ابزار و شبکه های جمع آوری داده ها
* درک و توصیف سیستم های زمینی از طریق تحقیق و تجزیه و تحلیل آن داده ها.
* ارزیابی و پیش بینی تغییرات این سیستم ها در طول زمان.
* مشارکت، مشاوره و اطلاع رسانی به سازمان های عمومی و شرکا با اطلاعات مهم
* مدیریت منابع برای بهبود جامعه، اقتصاد و محیط زیست

## لابراتوار NOAA Earth System Research

ESRL  یک لابراتوار جامع در نائو که لابراتورهای قدیم را در خود جمع کرده است.

این لابراتوار به چهار بخش1 نظارت جهانی، 2علوم فیزیکی، 3علوم شیمی و 4 سیستم های جهانی تقسیم می شود

از اهداف مهم این لابراتوار می توان به " **درک مکانیزم های جوی که آب و هوای زمین را کنترل می کند**. " اشاره کرد، اهداف بسیار مهم دیگری نیز برای این لابراتوار موجود می باشد که در راستای این پژوهش قرار ندارد .

## بازتحلیلی هواشناختی(طراحی سامانه درک و فهم داده های اقلیم)[[7]](#footnote-7)

بازنگری یا بازتحلیلی هواشناسی، درک و استخراج دانش از داده های گذشته اتمسفر زمین می باشد .

نمونه کارهایی که در دنیا با این رویه مطالعاتی را انجام داده اند در زیر بیان شده است.

* NCEP/NCAR Reanalysis
* JCDAS
* . ECMWF re-analysis

می توان به این سه مورد اشاره کرد که در ادامه هر یک به تفکیک توضیح داده خواهد شد.

# مدل NCEP/NCAR Reanalysis

مجموعه داده ایست که از سال 1948-تاکنون به طور مسمتر در حال بروز رسانی بوده است که بصورت گرید جهانی مجموعه داده ای را از وضعیت اتمسفر زمین فایل شده است. این مجموعه داده از تلفیق دو روش مشاهدات جوی(تجربی) و روش پیش بینی آب و هوا بصورت عددی درست گردیده است

این محصول یک محصول مشترک [[8]](#footnote-8)NCEP و NCAR[[9]](#footnote-9) می باشد

### Accessing the data

دسترسی به این داده ها به صورت آزاد می باشد که توسط NCEP و ESRL ارائه می گردد.

این داده ها در فایلهایی با فرمت های [Netcdf](https://en.wikipedia.org/wiki/Netcdf" \o "Netcdf) and [GRIB](https://en.wikipedia.org/wiki/GRIB) می باشد که برای کار با آن کتابخانه هایی تعبیه شده است.

## مدل CISL Research Data Archive

نوع دیگری از ارشیو داده های آب و هوا می باشد که برای پیش بینی استفاده می شود. این مدل نیز در NCAR بوجود امده است.

# مدل ECMWF re-analysis

نوع دیگری از ارشیو داده های آب و هوا می باشد که در دوره های مختلفی جمع آوری شده است.این دوره ها بصورت زیر رده بندی شده است:

|  |
| --- |
| Global Reanalyses |
| ERA5 (Jan 2008 - present) |
| CERA-20C (Jan 1901 - Dec 2010) |
| ERA-20C (Jan 1900 - Dec 2010) |
| ERA-Interim (Jan 1979 - present) |
| ERA-Interim/LAND (Jan 1979 - Dec 2010) |
| ERA-20CM (Jan 1900 - Dec 2010) Final |
| ERA-40 (Sep 1957 - Aug 2002) |
| ERA-15 (Jan 1979 - Dec 1993) |
| CERA-SAT (Jan 2008 - Dec 2016) New! Monthly datasets added |

پس از مطالعه نحوه ارائه داده ها به مشتریان، معلوم شد که می توان داده ها را در در فورمت استاندارد .grib و .nc بدست آورد که می بایست مورد پردازش محاسباتی قرار بگیرد تا بتوان داده ها را استفاده نمود.

# فصل اول تعریف مسئله

# **تعریف** مسئله

ذخیره و بازیابی و تحلیل داده های هوایی، کارایی بالایی در زمینه های مختلف تحلیل مهندسی دارد، از جمله در بحث های شبیه سازی مدل های فیزیکی که از سیستم های کامپیوتری نشأت میگیرد.

بنابراین مسئله پیش رو، ذخیره، داده کاوی و تحلیل آن برای ایجاد دانش که منجر به ایجاد اپلیکیشن های مختلف مهندسی می شود.

# فصل دوم دریافت و ذخیره داده ها از وبگاه weather.uwyo.edu

# دریافت و ذخیره داده ها

دادهای هواشناسی در برخی از سایتهای علمی قرار داده شده است. برای این مسئله از سایت زیر با پرس و جوی نقطه خاص مورد نظر، داده ها ذخیره می شود.

http://weather.uwyo.edu/cgi-bin/sounding?region=mideast&TYPE=TEXT:LIST&YEAR=2009&MONTH=1&FROM=all&To=0100&STNM=40800

این دانلود ها از طریق پیاده سازی برنامه که در زبان جاوا صورت گرفت، بصورت فایل .data ذخیره می گردد که تمام داده های یک ماه مشخص را دارا است.برخی از مشکلاتی که برای این الگوریتم ذخیره پیش آمد به صورت زیر آورده شده است.

* در مدل ابتدایی دانلود، مدت زمان ذخیره هر پرس وجو حدود 40 ثانیه زمان می برد.
* برای افزایش کارایی دانلود از نظر زمانی و کاهش این زمان نسبتا طولانی، مطالعه یک هفته ای صورت گرفت و چندین هفته برای پیاده سازی آن

پس از ذخیره داده های خام، نمونه ای از آن بصورت زیر آورده شده است:

|  |
| --- |
| **40800 OIFM Esfahan Observations at 12Z 01 Jan 2009**  **-----------------------------------------------------------------------------**  **PRES HGHT TEMP DWPT RELH MIXR DRCT SKNT THTA THTE THTV**  **hPa m C C % g/kg deg knot K K K**  **-----------------------------------------------------------------------------**  **1000.0 45**  **925.0 725**  **850.0 1443**  **839.0 1550 12.2 2.2 50 5.37 180 10 300.0 316.3 301.0**  **700.0 3038 3.2 -2.8 65 4.47 240 54 306.0 320.0 306.8**  **500.0 5680 -13.7 -19.7 61 1.61 245 58 316.3 321.8 316.6**  **400.0 7340 -25.9 -34.9 43 0.50 245 91 321.2 323.1 321.3**  **300.0 9370 -39.1 -56.1 15 0.06 250 111 330.1 330.4 330.2**  **250.0 10610 -45.3 260 143 338.6 338.6** |

**از داده های فوق می توان دریافت که**

* داده ها بصورت متن خام بدست آمده اند.
* داده ها در برخی جاها بدون مقدار می باشند و باید بعنوان پوچ در نظر گرفته شود.
* محل قرارگیری داده ها در ستون فایل برای هر فیلد تقریبا ثابت است.
* در هر بار پرسوجو متن، تمام روز های یک ماه ذخیره می شود، که می بایست بصورت درختی و نشانه دار ذخیره گردد، تا در ادامه توانایی طبقه بندی کردن داشته باشد.
* داده ها از سال 1973 میلادی تا به امروز در این وبسایت قرار داده شده است، که بصورت مناطق مختلفی از جهان قابل دسترسی است.
* بسیاری از روزهای سال بدون داده می باشد، با این وجود به دلیل زیادی سال های متوالی، می توان گفت که پایگاه داده خوبی قابل جمع آوری است.

با برسی داده های بدست آمده می توان پارامتر های فیزیکی زیر را دسته بندی کرد.

|  |  |  |
| --- | --- | --- |
| Parameter | Description | Units |
| PRES | [Atmospheric Pressure](http://glossary.ametsoc.org/wiki/Atmospheric_pressure) | [[hPa]](http://glossary.ametsoc.org/wiki/Hectopascal) |
| HGHT | [Geopotential Height](http://glossary.ametsoc.org/wiki/Geopotential_height) | [[meter]](http://glossary.ametsoc.org/wiki/Meter) |
| TEMP | [Temperature](http://glossary.ametsoc.org/wiki/Temperature) | [[celsius]](http://glossary.ametsoc.org/wiki/Celsius_temperature_scale) |
| DWPT | [Dewpoint Temperature](http://glossary.ametsoc.org/wiki/Dewpoint) | [[celsius]](http://glossary.ametsoc.org/wiki/Celsius_temperature_scale) |
| FRPT | [Frost Point Temperature](http://glossary.ametsoc.org/wiki/Frost_point) | [[celsius]](http://glossary.ametsoc.org/wiki/Celsius_temperature_scale) |
| RELH | [Relative Humidity](http://glossary.ametsoc.org/wiki/Relative_humidity) | [%] |
| RELI | [Relative Humidity with respect to Ice](http://glossary.ametsoc.org/wiki/Relative_humidity_with_respect_to_ice) | [%] |
| MIXR | [Mixing Ratio](http://glossary.ametsoc.org/wiki/Mixing_ratio) | [gram/kilogram] |
| DRCT | [Wind Direction](http://glossary.ametsoc.org/wiki/Wind_direction) | [[degrees true]](http://glossary.ametsoc.org/wiki/Azimuth) |
| SKNT | [Wind Speed](http://glossary.ametsoc.org/wiki/Wind_speed) | [[knot]](http://glossary.ametsoc.org/wiki/Knot) |
| THTA | [Potential Temperature](http://glossary.ametsoc.org/wiki/Potential_temperature) | [[kelvin]](http://glossary.ametsoc.org/wiki/Kelvin_temperature_scale) |
| THTE | [Equivalent Potential Temperature](http://glossary.ametsoc.org/wiki/Equivalent_potential_temperature) | [[kelvin]](http://glossary.ametsoc.org/wiki/Kelvin_temperature_scale) |
| THTV | [Virtual Potential Temperature](http://glossary.ametsoc.org/wiki/Virtual_potential_temperature) | [[kelvin]](http://glossary.ametsoc.org/wiki/Kelvin_temperature_scale) |

که در ادامه به تعریف هر یک از این پارامترها آورده شده است.

## تعریف پارامترهای فیزیکی

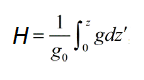
فشار اتمسفرPRES ))

فشار اعمالی ناشی از وجود اتمسر در اطراف زمین

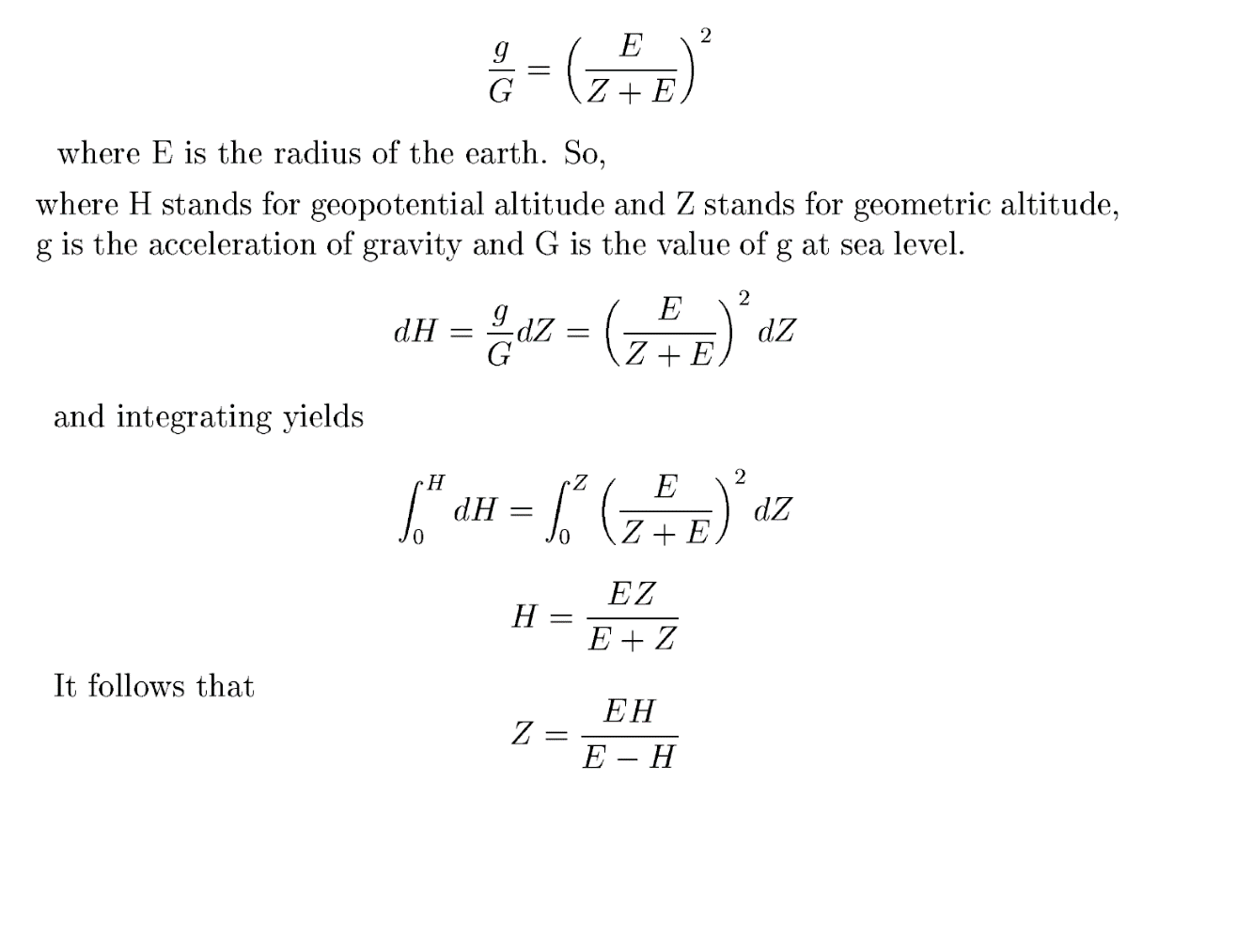
**ارتفاع پتانسیل جغرافیایي – (HGHT)**

ارتفاع یک نقطه مورد نظر در اتمسفر در واحد انرژی پتانسیل گرانشی برای یک ارتفاع مورد نظر نسبت

به سطح ازاد دریاها .که به صورت مفهوم ریاضی رابطه زیر می باشد.



از معادلات نیوتن حاکم بر جهان می توان رابطه انتگرالی بالا را به صورت زیر ساده سازی کرد.

 مشاهده می شود که مقدار ارتفاع اصلی ِ از سطح دریا(Z) به دست می آید.

**دما (TEMP )**

دما کمیتی که در این مدل با دماسنج در هر نقطه جغرافیایی اندازهگیری می شود..

**دمای نقطه شبنم (DWPT)**

دمایی که هوای مخلوط با آب در فشاز ثابت و در مقدار بخار ثابت به منظور رسیدن به حالتی که

بخار آن هوا اشباع شود.

**رطوبت نسبي(MIXR)**

فشار جزئی بخار هوا نسبت به فشار اشباع بخار موجود در دمای بخار حاظر

**نسبت مخلوط (MIXR)**

مفدار جرم بخار هوا به جرم هوای خشک آن

**جهت وزش باد (DRCT)**

به صورت قرارداد و استاندارد مقدار زاویه بردار وزش باد نسبت به شمال در جهت عقربه ساعت.

بطور مثال بادی که از شرق می وزد 90 درجه ،باد غرب 270 )یا - 90 (می باشد.) باد جنوب 180 درجه)

**سرعت وزش باد (SKNT)**

مقدار سرعت وزش باد که جهت آن در پارامتر فوق بررسی شد

**دمای پتانسیل (THTA)**

دمایی که یک بخش از هوای اشباع نشده از هوای خشک دارا است بشرطی که از حالت اولیه خود به صورت آدیاباتیک و برگشت پذیر به فشار استاندارد (kp100)که بصورت ریاضی رابظه زیر می شود

ams2001glos-Pe37

κ is the Poisson constant

# فصل سوم تحلیل داده ها

## تحلیل داده ها

پس از تعریف پارامتر های فیزیکی در فصل گذشته، کاوش بروی مقادیر این پارامتر ها، می تواند منجر به تولید دانش شود. به عبارتی با تلفیق منطق و الگوریتم های ساده و قوی ریاضی می توان به عبارتی

یادگیری ماشین انجام داد .دو دیدگاه حاکم بر داده کاویِ بصورت نمودارهای رگرسیون اماری،

* بایاس بودن
* واریانس بالا،

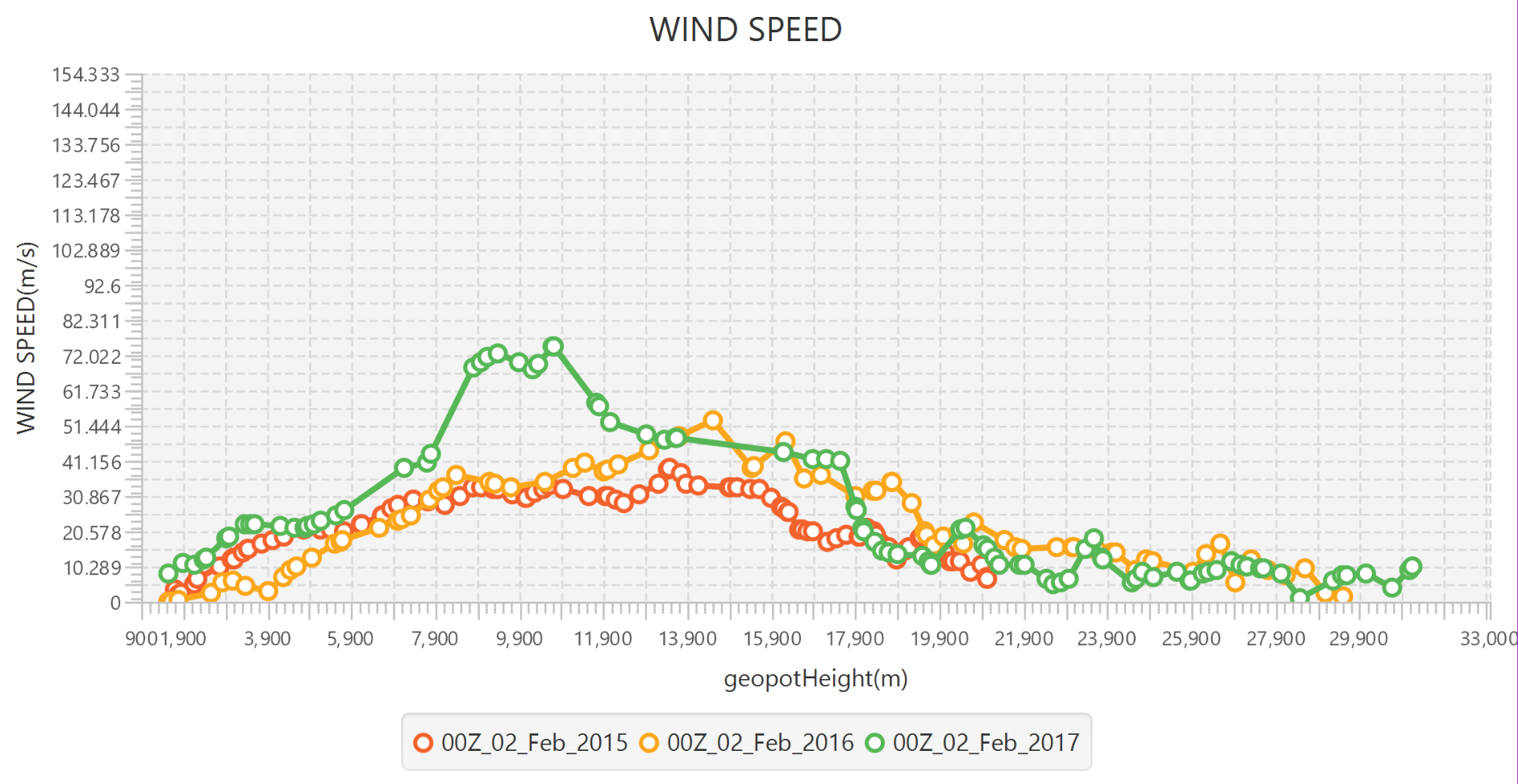
منجر به پیش بینی با بهترین احتمال صحت نتایج را فراهم می کند این مفاهیم در نمودارهای شکل زیر قابل مشاهده است..



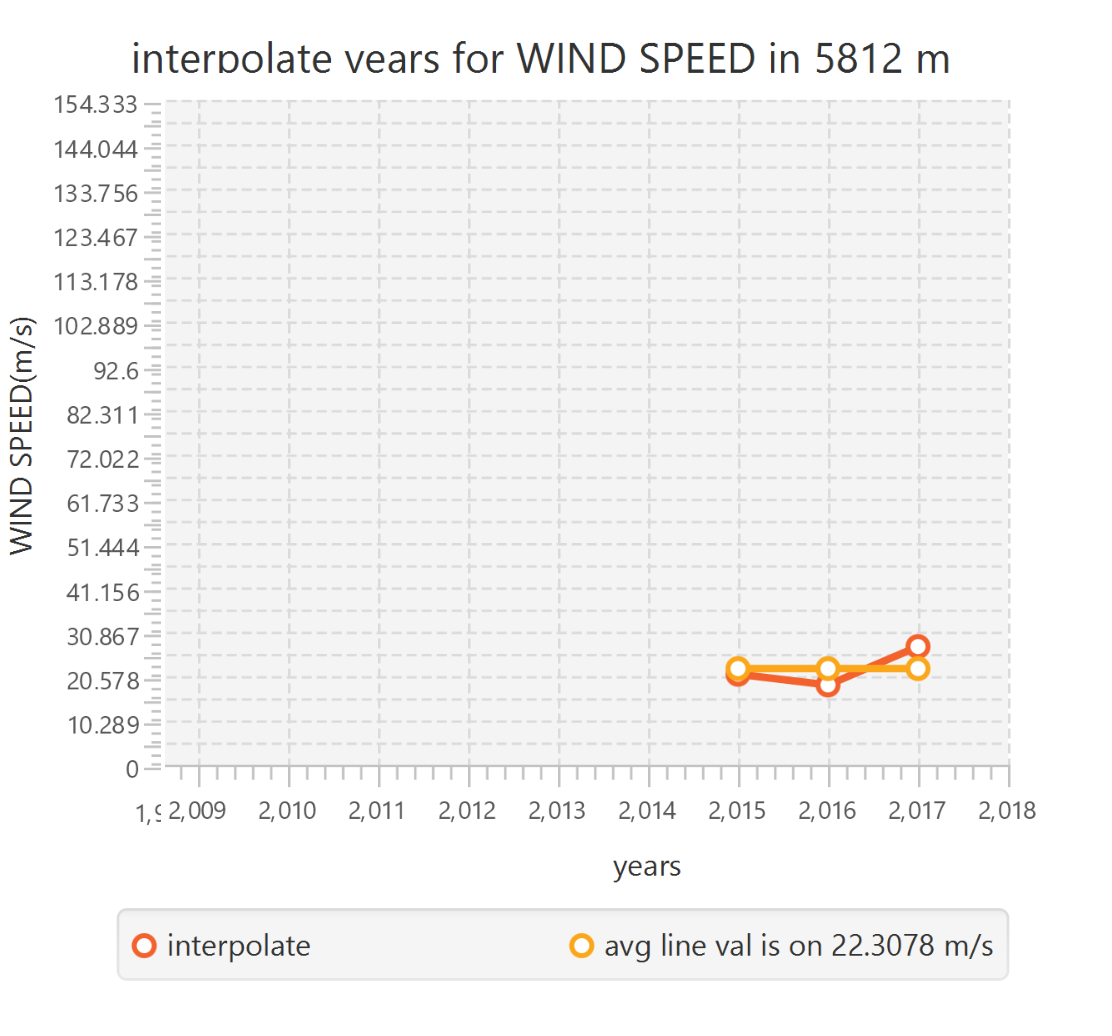
شکل 5 نمودارهای برازش داده ها بر مبنای نوع مدل سازی

## ارزیابی از داده ها

برای اینکه اهمیت داده ها درست بیان شود، نمونه ای از مقدار پارامتر های فیزیکی در آورده شده است.که در شکل زیر قابل مشاهده است.

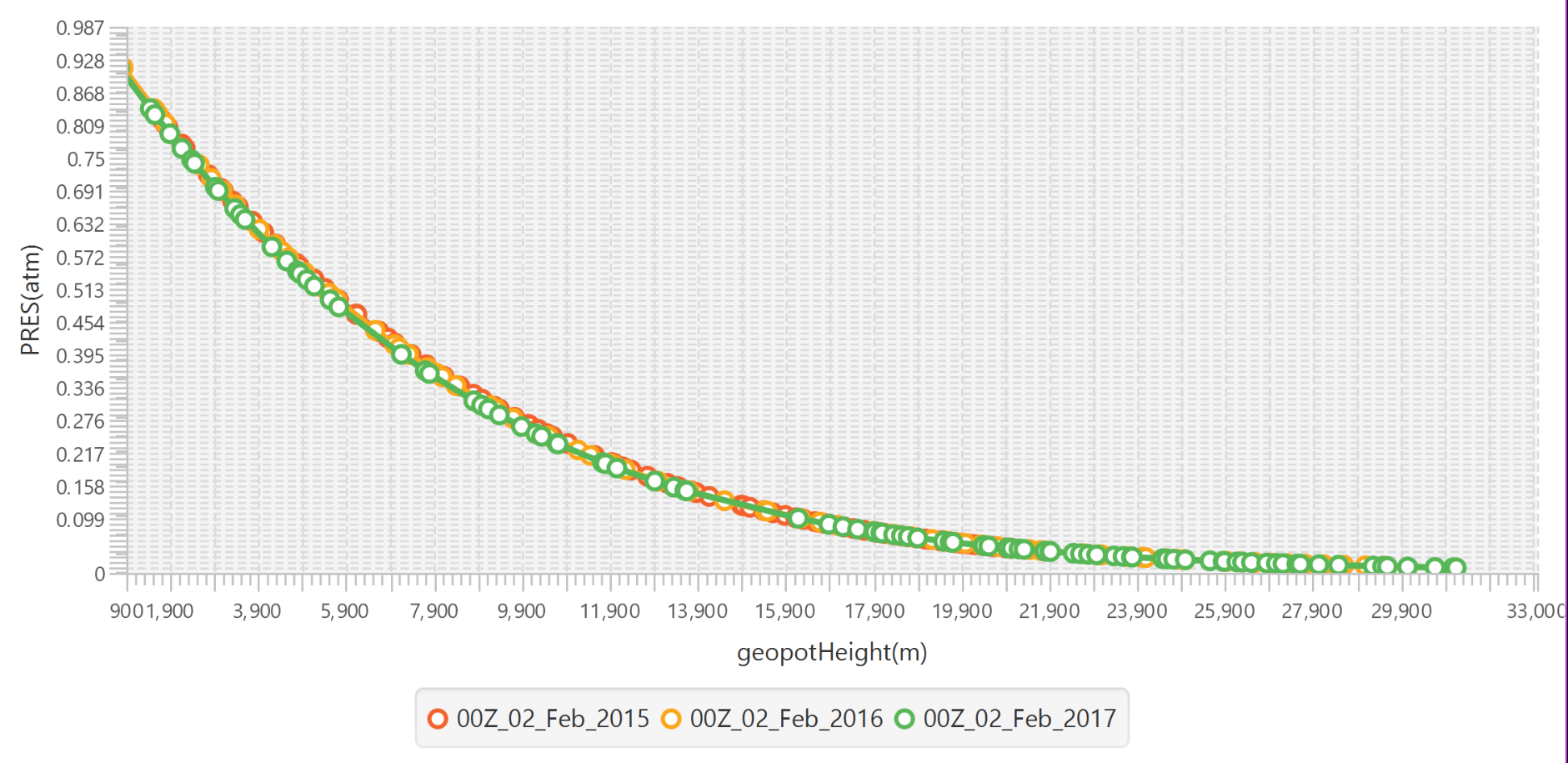


شکل 6داده های سرعت باد در ارتفاع های مختلف



شکل 7 میان یابی سرعت باد، از سالهای مختلف در ارتفاع مورد نظر

. از شکل فوق می توان مشاهده کرد که براورد قابل قبولی(واریانس میانی در حدود صفر می باشد) .می توان پارامترهای دیگر را مانند شکل 3، رسم نمود.



شکل 8 تغییرات فشار جو بر حسب ارتفاع آن

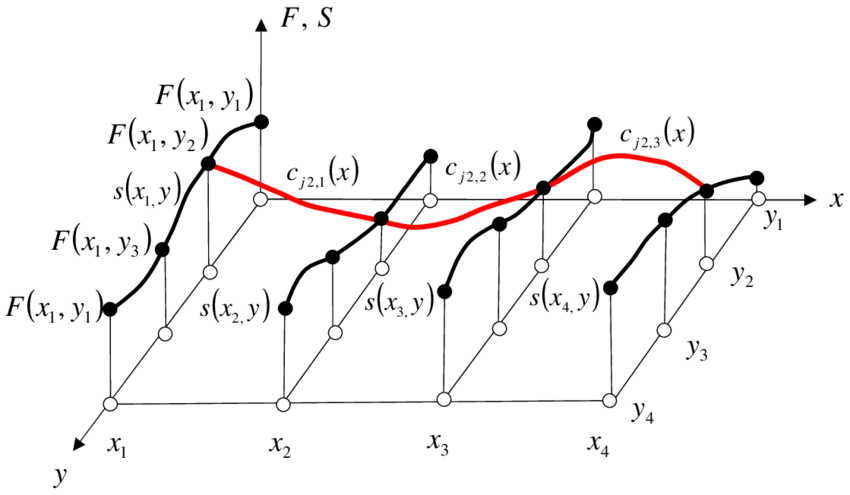
## میان یابی درارتفاع موردنظر در توابع اتمسفر

برای اینکه به توان در نقطه ارتفاع های مطلوب مقداری را بدست آورد نیازمند برون-درون یابی از داده های موجود در ارتفاع های مختلف است.

روش هایی که برای این مدل استفاده شده است این موارد می باشد:

* روش میانیابی خطی بین دو نقطه .
* روش مربعی مرتبه دوم
* روش انطباق مکعبات مرتبه سوم

هر سه روش مقادیر معقولی را با دقت مناسبی(به لحاظ کمی برای هر مشخصه فیزیکی اتمسفر این دقت متفاوت می باشد ) پیش بینی می کنند.



شکل مفهوم روش میان یابی 9Bi-cubic spline interpolation

## شبکه های عصبی بر روی داده ها

به دلیل وسعت بسیار دنیا، و محدودیت در تعداد ایستگاه هواشناسی و هم چنین محدود در تعداد داده برداری در ارتفاع های مختلف(شکل 4)، نیاز به پیش بینی مقدار پارامترهای فیزیکی در جاهایی که ایستگاه هواشناسی و ارتفاع خاص وجود ندارد، ضروری می باشد. یکی از راه حل ها، و الگوریتم های محاسباتی، شبکه های عصبی می باشد که پیشرفت چشم گیری داشته است و از قدرت خوبی برای جواب دهی، برخوردار است.



شکل 10 ایستگاه های هواشناسی به عنوان نورون در شبکه عصبی محسوب می شود

پس برای براورد مقادیر پارامترهای فیزیکی در کل دنیا می توان شبکه عصبی ای را آموزش داد .به تبع مطالعه چگونگی مدل سازی این داده ها(تعداد نورون ها، تعداد لایه ها، تعداد خروجی و ورودی)دارای اهمیت بسیار می باشد.

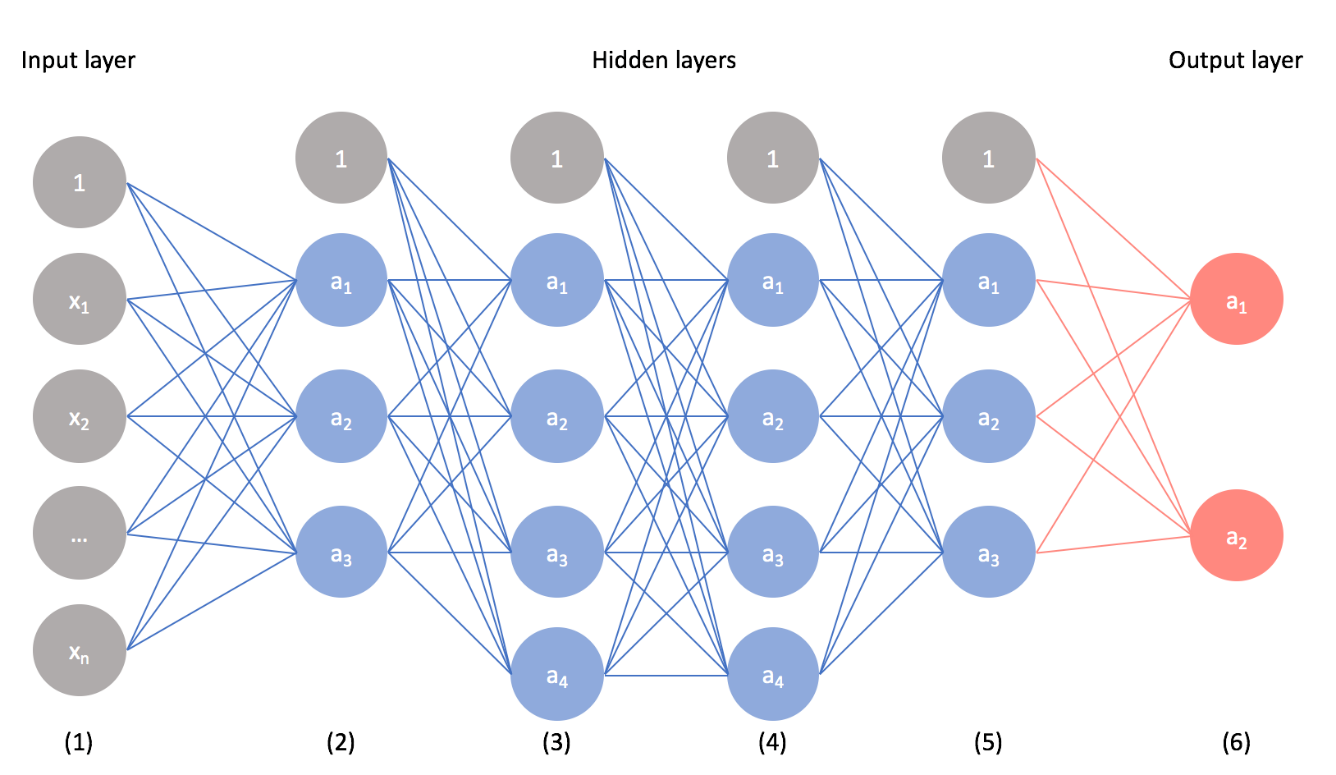
پس از پیاده سازی شبکه عصبی، نرم افزار قادر خواهد بود

* با دادن هر نقطه ای از جهان بصورت ورودی دو پارامتری طول و عرض چغرافیایی ،تمامی پارامترهای فیزیکی آن نقطه را با تقریب معلومی ، بدست آورد

## تشریح شبکه عصبی و الگوریتم پیاده سازی آن

شبکه های عصبی از روی نرون های عصب مغز الهام شده است، بگونه ای که سیناپس دنریت سیگنالی را دریافت می کنند روی آن اثر گذاشته و به عنوان خروجی از نرون، محسوب می شود .

هز شبکه عصبی از مجموعه ای لایه ای از نرون ها تشکیل شده است، هر نرون یک تابع می باشد که ورودی ای را دریافت می کند، بر روی آن وزن و ارزش دهی می دهد و با یک ترشولدی جمع می شود و به عنوان خروجی نرون محسوب می شود این خروجی می تواند به عنوان ورودی نرون دیگر باشد یا به عنوان خورجی کل مسئله نقش بگیرد در شکل زیر این مفهوم قابل مشاهده است.



شکل 11 مفهوم شبکه عصبی : حضور نرون ها در کنار هم

همان طور که از روی شکل قابل مشاده است شبکه عصبی به سه لایه اصلی تقسیم می شود.

* لایه ورودی
* لایه پنهان یا داخلی
* لایه خروجی

لایه ورودی در شبکه های عصبی

این لایه همان ورودی های مسئله می باشد، یا بطور ریاضی تر ورودی های تابع مورد نظر می باشد که می تواند تک ورودی یا چند پارامتره، بسته به نوع مسئله انتخاب شود،

لایه پنهانی شبکه های عصبی

این لایه دارای نرون هایی می باشد که سعی بر آموزش از روی داده های اموزشی دارند، خروجی نرونهای ورودی به عنوان ورودی این نرون ها در نظر گرفته می شود

لایه خروجی شبکه های عصبی

این لایه خروجی تابع مورد نظر را در بر دارد و به عنوان اخرین لایه در شبکه دیده می شود، تعداد پارامتر خروجی این لایه با تعداد خروجی تابع مسئله می بایست یکی در نظر گرفته شود.

## تابع آموزش شبکه

هر یک از نرون ها نیاز به یک تابع یادگیری دارند که در این روش بدلیل اینکه هر یک از داده ها اهمیت به سزایی دارند، انطباق کامل بر روی این داده ها اهمیت فراوانی دارد.

به همین دلیل از الگوریتم لونبرگ-مارکارد برای این یادگیری استفاده شده است.

## الگوریتم لونبرگ-مارکارد

لگوریتم لونبرگ-مارکارد روشی است برای یافتن کمینه یک تابع غیر خطی چند متغیره که به عنوان یک روش استاندارد برای حل مسئله کمینه مربعات برای توابع غیرخطی درآمده است.

الگوریتم لونبرگ-مارکارد(LMA)بین الگوریتم گاوس-نیوتون(GNA) و روش نزول گرادیانی درونیابی می‌کند. LMA از GNA مقاوم‌تر است، که یعنی در بسیاری مواقع، حتی اگر بسیار دورتر از کمینه نهایی شروع کرده باشد، جوابی را پیدا می‌کند. از دیگر سو، برای تابع‌های خوشرفتار و پارامترهای آغازین معقول، LMA کمی کندتر از GNA است. LMA پرطرفدارترین الگوریتم برازش خم است و کاربران کمی ممکن است به روش‌های دیگر برازش خم نیاز پیدا کنند.

مانند سایر الگوریتم‌های کمینه‌سازی عددی، الگوریتم لونبرگ-مارکارد یک رویه تکراری است. برای شروع کمینه‌سازی، کاربر باید یک حدس آغازین برای بردار pپارامترها ارائه کند. در بسیاری مواقع یک حدس ناآگاهانه استاندارد مانند {p^{T}}=(1,1,1,1) به خوبی کار می‌کند. در جاهای دیگر، الگوریتم تنها وقتی کار می‌کند که حدس آغازین تا حدی به جواب نهایی نزدیک باشد.

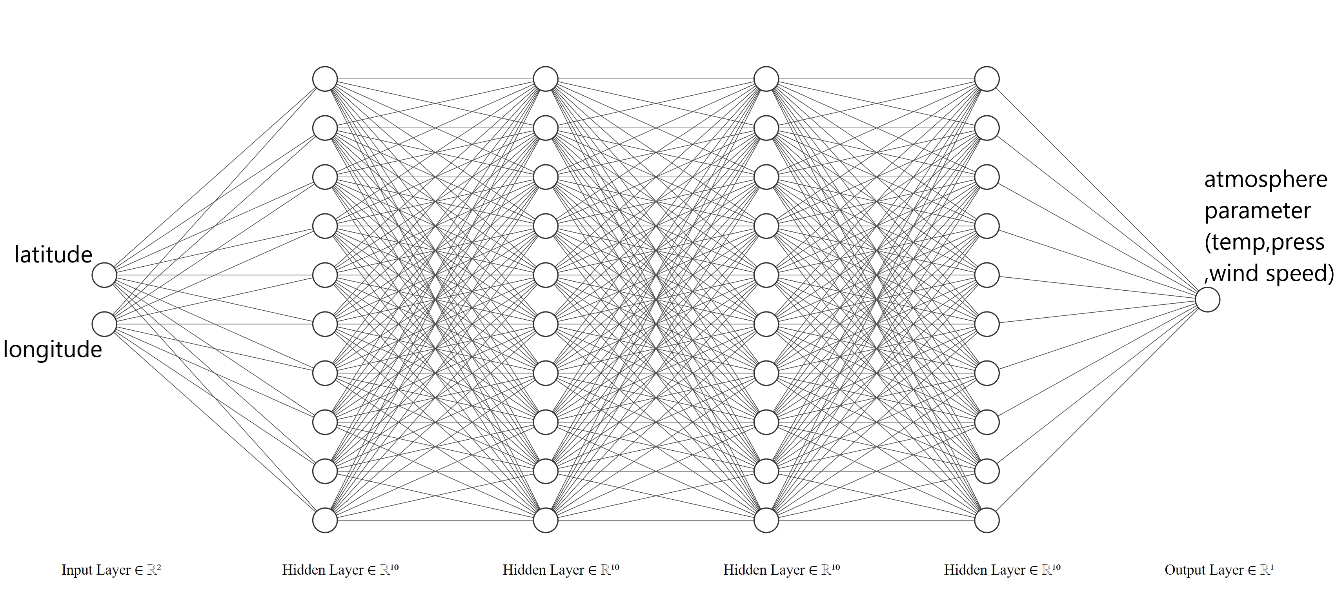
در هر گام تکرار، بردار {p}پارامترها با یک تخمین جدید {{p+q}} جایگزین می‌شود. برای دستیابی به {{q } } توابع f\_{i}({p+q} )} با خطی‌سازیشان

f(p+q) == f(p)+Jq

تخمین زده می‌شوند که {J} } ژاکوبین {f} } در {p} است

## پیاده سازی الگوریتم برای مسئله مورد نظر

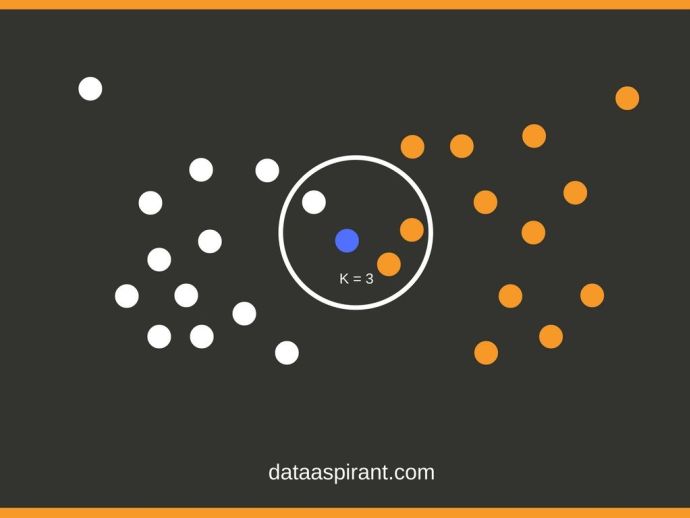
با استفاده از مفاهیم و روش های گفته شده در فوق شبکه عصبی این مسئله به صورت زیر خواهد بود.

چهار لایه پنهان با تعداد 10 نرون برای هر لایه برای مدل سازی فیزیک مسئله در نظر گرفته شده است، تا دقت قابل قبولی برای هر یک از پارامترهای اتمسفر بدست آید 

شکل 12 شبکه عصبی مورد استفاده در این مسئله، پارامتر های ورودی و خروجی مشخص شده است

## الگوریتم k همسایه نزدیک[[10]](#footnote-10)

این الگوریتم بسیار ساده و کاربردی می باشد، بگونه ای که برای بدست آوردن مقدار یک ورودی که داده آن موجود نیست ، از k همسایه موجود در اطراف خود کمک می گیرد،تا مقادیر خود را بدست آورد، شکل زیر گویای این مطلب می باشد.

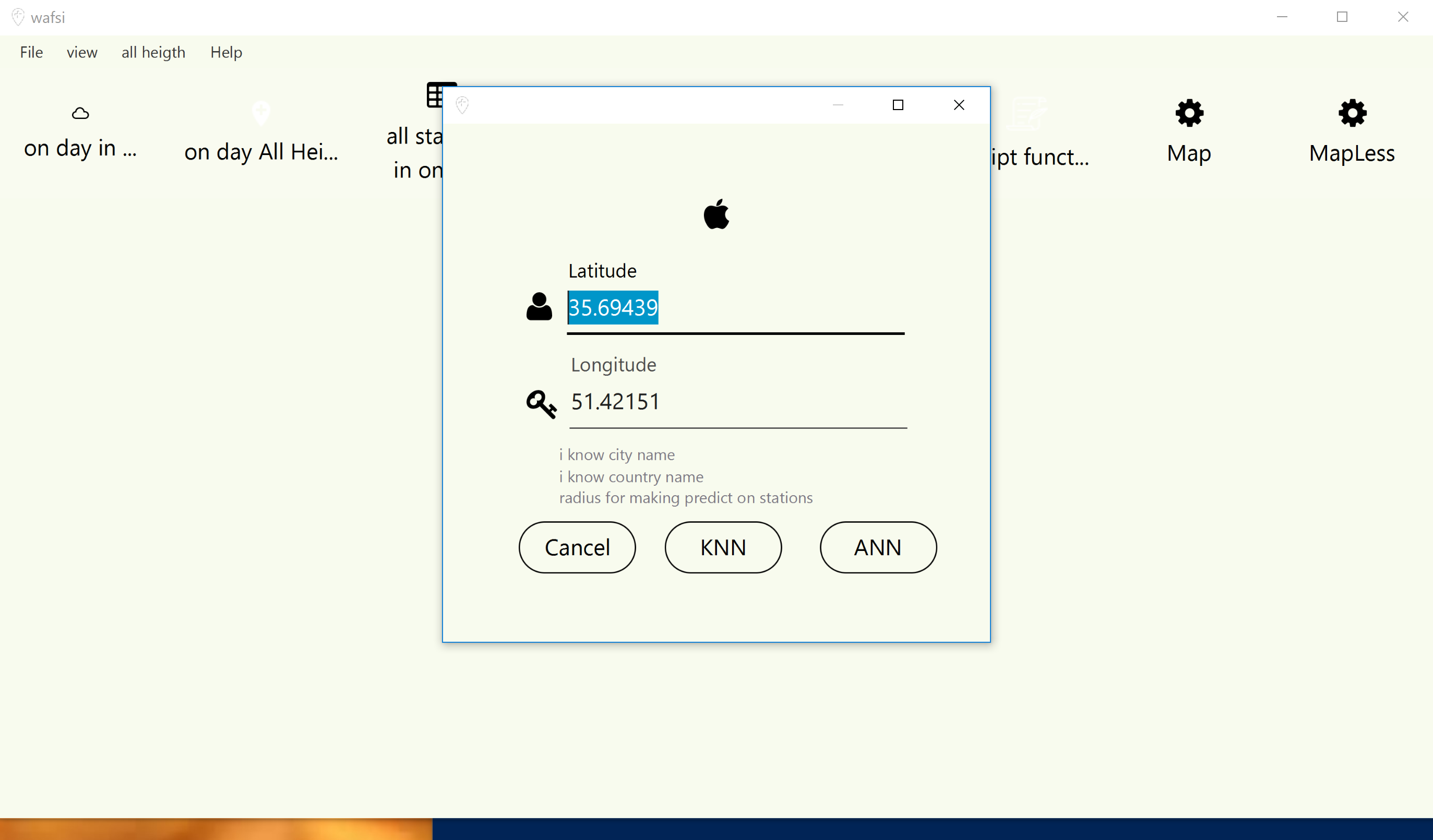


شکل 13 مفهوم الگوریتم knn

معیار همسایگی در مدل این مسئله فاصله ایستگاه های هواشناسی از یکدیگر است.که بصورت معکوس در رابطه ریاضی آورده می شود.

## پیاده سازی دو الگوریتم ذکرشده یادگیری ماشین

دو روش گفته شده پس از بررسی و جستجوی کتاب خانه ها، مطالعات مختلف و بررسی های



شکل 14 پیاده سازی دو الگوریتم knn و Ann برای یک نقطه دلخواه از زمین

به صورت نمونه مقدار سرعت باد در روز 4 آبان در یک نطقطه خاص در جدول زیر به نمایش گذاشته شده است.

|  |  |  |
| --- | --- | --- |
| KNN | ANN | (Lat ,Long) |
| 124.2320 km/h | km/h 124.14 | (35.69439, 51.42151) |

|  |  |
| --- | --- |
| =7.045e-4 | error |

در این مثال هر دو روش با دقت قابل قبولی مقادیری را ارائه نموده اند.

1. Weather-Climate connection science [↑](#footnote-ref-1)
2. National Oceanic and Atmospheric Administration [↑](#footnote-ref-2)
3. National Weather Service [↑](#footnote-ref-3)
4. Earth System Research Laboratory [↑](#footnote-ref-4)
5. atmospheric rivers [↑](#footnote-ref-5)
6. National Oceanic and Atmospheric Administration [↑](#footnote-ref-6)
7. meteorological reanalysis(Climate Data Assimilation System) [↑](#footnote-ref-7)
8. National Centers for Environmental Prediction [↑](#footnote-ref-8)
9. National Center for Atmospheric Research [↑](#footnote-ref-9)
10. KNN(k-nearest neighbors algorithm) [↑](#footnote-ref-10)