Models de xarxes neuronals versus  models lineals per la predicció de sèries temporals

ÍNDEX

1. Introducció

2. Motivació

3. Dades (anàlisi exploratori..)

Definicions?

4. Metodologia (quins mètodes, funcionament...)

5. Aplicació

6. Resultats

7. Conclusions

1. Introducció

2. Motivació

3. Dades

3.1 Presentació estadística

*- Descripció de les dades*

Per posar a prova l’estudi s’utilitzaràn dades sobre els moviments turístics en frontera. Aquestes dades tenen l’objectiu de proporcionar estimacions mensuals i anuals sobre l’entrada a espanya de visitants no residents a Espanya.

*- Origen de les dades*

Aquestes dades es recolleixen des del 1996 per l’Institut d’Estudis Turístics (Frontur) i a partir del 2015 l’institut Nacional d’Estadística s’encarrega de penjar-les.

*- Sistemes de classificació*

Per l’estudi, es decideix utilitzar les dades des del 2000 fins al 2019

relatives a:

* **Comunitats Autònomes**

Catalunya

Madrid

València

* **Tipus de Visitant**

Turistes

Excursionistes

* **Vies d’accés**

Aeroport

Carretera

Port

*- Unitat estadística*

La unitat de l’anàlisi són els viatgers, els visitants, els viatges i les excursions

4. Metodologia

Definim una sèrie temporal com una seqüència d’observacions (dades) sobre una variable ordenades cronològicament.

Un procés estocàstic és un conjunt de variables aleatòries associades a diferents instants del temps.

Una sèrie temporal és la realització d’un procés estocàstic.

En referència als mètodes clàssics ens centrarem en l’anàlisi estocàstic de les sèries temporals, per això, cal introduir els mètodes ARIMA.

4.1 Mètodes clàssics (ARIMA)

Box and Jenkins (1970) van desenvolupar models que tenen en compte la dependència entre les dades, és a dir, cada observació en un moment donat és modelada en funció dels valors anteriors.

Els models és coneixen amb el nom genèric d’ARIMA ( models Autoregressius Integrats de Mitjana Mòbil) que es poden separar en tres components AR (Autoregressiu), I (Integrat) i MA (Mitjana mòbil).

***Processos Autoregressius AR(p):***

Utilitzant l’operador de retard podem escriure aquests processos lineals com:

On

***Processos Mitjanes mòbils MA(q):***

Utilitzant l’operador retard:

On

***Processos mixtes ARMA (p,q) (la combinació d’AR i MA) :***

Utilitzant l’operador retard:

***Processos ARIMA (p,d,q)***

Si extenem un model ARMA per permetre que s’incloeixin arrels unitàries ens estem referint als models ARIMA d’ordre (p,d,q) on d és el nombre de vegades que s’aplica l’operador diferencia:

On

(Polinomi AR(p))

(Polinomi (MA)(q))

Els models ARIMA com s’ha mencionat anteriorment, són models estocàstics que utilitzen variacions i regressions de dades amb la finalitat de trobar “patrons” per així poder realitzar una predicció futura. Es tracta d’un mètode en que les estimacions futures venen explicades per les dades del passat i no per variables independents.

En els models ARIMA, els valors estimats es suposa que són una combinació lineal dels valors passats i dels errors passats.

Sovint, ens trobem amb series temporals que presenten un clar comportament estacional, per tractar aquestes sèries es poden utilitzar dos opcions:

1. Considerar oscil·lacions estacionals deterministes.

2. Considerar oscil·lacions estacionals estocàstiques.

Com ja s’ha mencionat en aquest estudi es treballarà amb un anàlisi estocàstic, per tant, considerarem el segon tipus d’oscil·lacions. Per fer-ho definirem una “s” com el número d’observacions incloses en el cicle estacional complet (en el nostre cas s=12 ja que tenim dades mensuals)

Quan es tenen aquest tipus de models estacionals estem parlant de processos SARIMA.

***Models estacionals SARIMA***

Aquests models permeten captar de manera simultània la component “regular” (dependència respecte dels valors adjacents) i la component estacional (dependència respecte de les mateixes estacions d’anys diferents). Aquest model es denomina com: SARIMA(

On els primers tres paràmetres fan referència a la part regular i els tres següents a la part estacional.

On:

(Polinomi AR(p))

(Polinomi MA(q))

(Operador diferència estacional)

4.1.1 Modelització manual vs automática

Dins dels models ARIMA, distingim entre dos maneres de modelitzar; una manera automàtica (auto.arima) i una manera més manual:

* Modelització manual:

En aquest cas, cal que nosaltres ens encarreguem d’estudiar la sèrie, determinar si necessita ajusts i decidir el model adequat per fer prediccions.

* Modelització automàtica (auto.arima())

Aquest mètode s’encarrega de seguir els mateixos passos que l’anterior però de manera automàtica, és a dir, l’R s’encarrega de crear un model idoni per les nostres dades.

Aquesta funció es troba dins del paquet forecast() i bàsicament escolleix el model que té un valor del AIC o BIC més baix, i per tant, el millor model.

Tot i que aquesta funció sembla que és solució fàcil i ràpida per la predicció de sèries temporals i per tant no caldria la modelització manual s’ha de tenir en compte que tal i com diu el nom és automàtica, amb això el que es vol dir és que és una funció que té uns paràmetres establerts i hi ha vegades que pot ser que les dades no expressin clarament el seu comportament i per tant la funció no sigui capaç de triar els paràmetres correctament.

4.1.2 Box and Jenkins aplicat al turisme

Una vegada definits els models, Box and Jenkins proposa una sèrie de passos clau que s’han de seguir per tal de modelar correctament les nostres dades i poder fer una predicció:

Definició del problema i dels objectius

Identificació

Estimació

El model és vàlid?

Utilització del model

S’explicarà una de les series tractades amb profunditat per tal d’ensenyar el procés que s’ha realitzat per cada una de les sèries.

La sèrie explicada serà el nombre de total de turistes per Comunitat Autònoma.

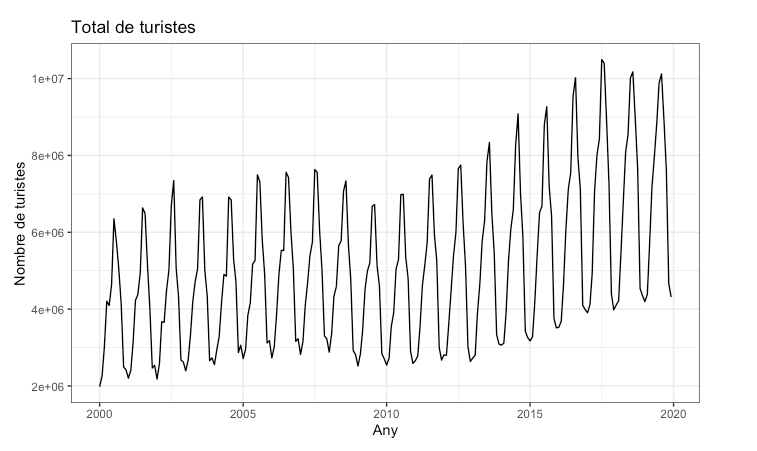
***1. Identificació:*** L’objectiu és identificar un model que de manera versemblant hagi pogut generar els valors observats:

Primer de començar, cal recordar que necessitem que la nostra sèrie temporal compleixi les condicions d’estacionarietat i ergodicitat.

Diem que una sèrie temporal és estacionària quan la mitjana i la variabilitat es mantenen constants al llarg del temps, és a dir, no es en funció del temps i a més a més, no presenta tendència.

Un procés estocàstic és ergòdic si es possible estimar de manera consistent les seves característiques a partir d’una realització seva.

El primer pas per la identificació és crear un gràfic que ens mostri el comportament de la sèrie, també es realitzen els gràfics de la FAS (Funció d’Autocorrelació Simple) i la FAP (Funció d’Autocorrelació Parcial) per determinar si són necessàries les transformacions:



Si ens fixem en el comportament de la sèrie des del 2000 fins l’actualitat, el primer que veiem és que la sèrie es comporta de manera molt similar durant tot el període, és a dir, es registren pujades i baixades similars al llarg dels anys en les diferents estacions de l’any (estacionalitat). Veiem que això té sentit, ja que es coneix que hi ha certs mesos de l’any que el nombre de turistes serà o no major.

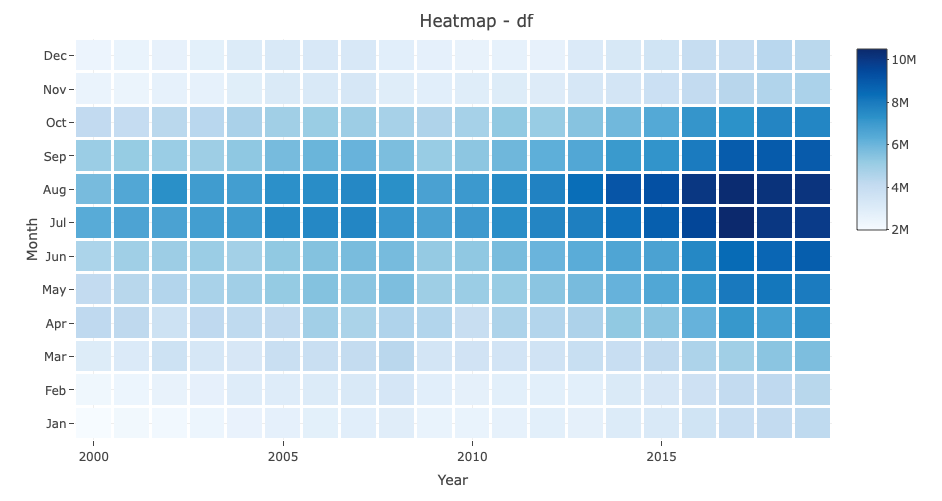
A més a més, podem veure que hi ha hagut un creixement del nombre de turistes amb el pas dels anys (tendència). Aquest augment es podria ser a causa de les millores en les infraestructures de comunicació i el desenvolupament de les noves tecnologies. Tot i així, veiem que dins d’aquest creixement també existeix una certa variació, per exemple, podem veure que hi ha una baixada del nombre de turistes durant els anys 2008-2012 que si ens posem en context veiem que durant aquells anys es va produir una crisi financera global que va afectar a nivell mundial i per tant, el nombre de turistes es va reduir.

Abans d’estudiar si necessitem o no transformacions a les nostres dades, realitzarem un estudi de l’estacionalitat de la nostra sèrie.

Imagen que contiene texto, mapa

Descripción generada automáticamente

D’aquest gràfic podem destacar sobretot que el turisme a Espanya sobretot augmenta a l’estiu, ja que es quan la majoria de treballadors tenen vacances, a més a més veiem que a mesura que el bon temps augmenta el nombre de turistes també augmenta. A més a més, com s’ha mencionat abans veiem un increment en el nombre de turistes des del 2000 fins l’actualitat.



Aquest mapa de calor ens mostra també que durant tot el període estudiat, als mesos d’estiu és quan el nombre de turistes augmenta. El registre més alt el trobem al juliol del 2017 amb 10.493.438 turistes.

Imagen que contiene texto, mapa

Descripción generada automáticamente

Aquest gràfic ens mostra la visualització de les sèries amb el seus retards i serveix per identificar la correlació entre la sèrie i els retards. Veiem que del primer retard al setè, la relació entre la sèrie i el retard va perdent la seva forma lineal, a partir d’aquest fins l’últim la forma lineal va tornant fins que a l’últim hi podem veure dibuixada una línia quasi perfecta.

Amb això el que es vol dir és que s’observa una forta correlació entre el mateix mes al llarg dels anys (estacionalitat) i també entre els mesos propers.

Ara, procedirem a realitzar els gràfics d’autocorrelacions per veure quines diferències hem d’aplicar per tal d’aconseguir la condició d’estacionarietat en la nostra sèrie.

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Només mirant el gràfic de la Funció d’Autocorrelació Simple veiem, com hem pogut intuir amb el primer gràfic que les correlacions és comporten de manera molt similar al llarg dels anys, això ens fa veure que el turisme d’un any està correlacionat amb el turisme dels anys anteriors.

Al principi havíem vist que podia ser que la sèrie tingués una certa tendència creixent i ara mirant el gràfic ACF veiem que les correlacions disminueixen molt a poc a poc, per tant, decidim realitzar una diferenciació regular.

En aquest cas, només amb mirar el gràfic de la Funció d’Autocorrelació Simple veuríem que cal aplicar una diferència estacional i una diferència regular a la nostra sèrie temporal per tal de que es compleixi la condició d’estacionarietat.

Primer de tot aplicarem la diferència regular:

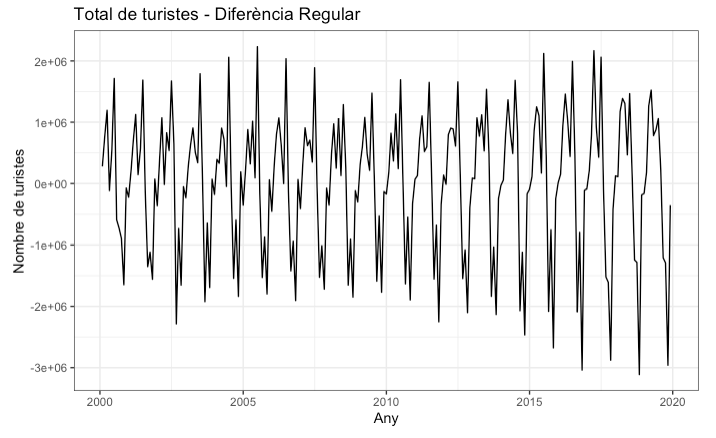


Imagen que contiene captura de pantalla

Descripción generada automáticamente

Podem veure que hem eliminat la tendència de la nostra sèrie però encara hi veiem un factor estacional molt present.

Imagen que contiene objeto

Descripción generada automáticamente

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Una vegada aplicada aquesta diferència podem veure que el factor estacional ha desaparegut i que les dades es mantenen mes o menys estables amb una mitjana de zero. Cal destacar que hi ha una forta baixada entre 2008-2010 que com s’ha mencionat abans aquesta baixada és deguda a la crisi global que hi va haver entre els anys 2008-2012.

A més a més aquí ja podríem veure que sembla que les dades compleixen la condició d’estacionarietat.

***2. Estimació:***

Com s’ha mencionat anteriorment, s’utilitzaran dos mètodes per la modelització de la sèrie temporal, el mètode manual i el mètode automàtic que amb l’R s’utilitza la funció auto.arima().

Per evitar influenciar-se s’utilitzarà primer el mètode manual.

* Mètode manual:

Abans de modelitzar ens hem de parar a pensar quin tipus de model volem estimar, com que hem aplicat una diferència estacional, sabem que estem parlant d’un model SARIMA( (En totes les sèries temporals tractarem aquests models) i per tant hem de decidir els seus paràmetres.

Com ja s’ha mencionat, els models SARIMA es separen en la part regular i la part estacional.

Per la part estacional mirarem els últims gràfics de la FAS i la FAP i per poder determinar el paràmetre de la part regular, el que es pot fer és reduir el número de retards que volem veure en els gràfics i per tant “apropar-nos” per tal de mirar la dependència respecte dels valors adjacents. S’ha utilitzat un retard de 12:

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Una vegada analitzats aquests gràfics es decideix escollir un model

SARIMA(2,1,1)(2,1,1) [12] i obtenim els següents resultats:

Coefficients:

ar1 ar2 ma1 sar1 sar2 sma1

-0.9205 -0.3904 0.3268 -0.9064 -0.3499 0.6533

s.e. 0.1953 0.1002 0.2089 0.1139 0.0709 0.1078

sigma^2 estimated as 5.378e+10: log likelihood = -3128.17, aic = 6270.34

* Mètode automàtic:

Per aquest mètode únicament ens cal la utilització de la funció auto.arima(). Si apliquem aquesta funció obtenim que el millor model és un SARIMA(0,1,1)(2,1,1) [12] i obtenim la següent informació:

Coefficients:

ma1 sar1 sar2 sma1

-0.6435 -0.8734 -0.3312 0.6450

s.e. 0.0630 0.1174 0.0706 0.1119

sigma^2 estimated as 5.516e+10: log likelihood=-3128.94

AIC=6267.89 AICc=6268.16 BIC=6285.01

***3. El model és vàlid?***

Una vegada hem decidit els paràmetres del nostre model, cal veure si aquests paràmetres són vàlids, per fer-ho hem de comprovar un seguit de condicions:

*1. Comprovar que les dades es distribueixen segons una distribució normal.*

Cal estudiar la normalitat dels residus mitjançant el test de Jarque Bera

*2. Tots els coeficients han de ser estadísticament significatius:*

Per això ens centrarem en mirar el valor-p de la significació assimptòtica dels coeficients seguint una distribució normal amb mitjana zero i variància 1.

*3. El model ha de complir amb les condicions d’estacionarietat i invertibilitat*

Cal comprovar que les arrels de la part AR i de la part MA estan fora del cercle d’unitat

*4. Els residus de l’estimació s’han de comportar segons soroll blanc.*

Per comprovar aquesta condició utilitzarem la funció d’R que ens retorna gràficament els residus escalats segons l'estimació de la seva variació (individual) i utilitzen la versió Ljung – Box del test de portmanteau.

Model 1

Imagen que contiene captura de pantalla

Descripción generada automáticamenteImagen que contiene captura de pantalla

Descripción generada automáticamente

Model 2

Imagen que contiene captura de pantalla

Descripción generada automáticamenteImagen que contiene captura de pantalla

Descripción generada automáticamente

*5. El model ha de ser estable*

***4. Prediccions i errors***

Una vegada validat model és l’hora de posar-lo a prova. En aquest punt realitzarem una predicció del model a temps real i estudiarem la capacitat predictiva.

Com s’ha mencionat anteriorment, es realitzaran les prediccions de gener a març de l’any 2020. L’estudi comença al gener de 2020 i cada principi de mes l’INE treu les dades de dos mesos anteriors, per tant, les prediccions es faran a temps real (Cada més s’actualitzaran les prediccions. A l’apartat de resultats es troben totes les prediccions i càlcul de la capacitat predictiva de cada sèrie i de cada un dels mètodes.

El 3 de febrer de 2020 l’INE publica les dades del mes de desembre de 2019 i per tant ja es pot realitzar la predicció del mes de gener, febrer i març:

Mètode Manual - SARIMA(2,1,1)(2,1,1) [12] – AIC: 6270,34

|  |  |  |  |
| --- | --- | --- | --- |
|  | gENER | fEBRER | Març |
| Gener | 4.243.462 | 4.442.793 | 5.573.931 |

Imagen que contiene objeto

Descripción generada automáticamente

Mètode Automàtic (auto.arima) - SARIMA(0,1,1)(2,1,1) [12] – AIC: 6267.89

|  |  |  |  |
| --- | --- | --- | --- |
|  | gENER | fEBRER | Març |
| Gener | 4.219.564 | 4.415.620 | 5.581.896 |

Imagen que contiene objeto

Descripción generada automáticamente

Podem observar que els resultats obtinguts pels dos mètodes són bastant similars, si ens fixem en el AIC veiem que és menor el valor del auto.arima() i això el faria un millor model.

A principis de març surten les dades del mes de gener de 2020 i per tant ja podem calcular la predicció de febrer i març i a més a més podem estudiar la capacitat predictiva de les prediccions pel mes de gener.

Primer de tot calcularem els errors de les prediccions del mes de gener comparant amb els valors reals. D’aquesta manera, podrem veure com d’acertat ha estat o no el nostre model.

|  |  |  |  |
| --- | --- | --- | --- |
|  | valor real | mètode manual | auto.arima |
| Predicció | 4.138.371 | 4.243.462 | 4.219.564 |
| error |  | -105.091 | -81.193 |

El model automàtic és el que la seva predicció s’apropa més al valor real ja que es la que té l’error més baix.

- PREDICCIONS FEBRER

4.2 Mètodes de Xarxes Neuronals

Les Xarxes Neuronals són un camp molt important dins de la Intel·ligència artificial. Aquestes Xarxes s’inspiren en el comportament conegut del cervell humà, referint-se a les seves neurones i les seves connexions.

Les Xarxes Neuronals tenen l’objectiu de crear models artificials que solucionin problemes difícils de resoldre mitjançant tècniques algorítmiques convencionals.

La forma bàsica d’una Xarxa Neuronal és:

Regla de propagació

Funció d’activació

Funció de sortida

Input

Output

Neurona

Capa de sortida

Capes ocultes

Capa d’entrada

Una xarxa neuronal té els següents components:

- *Capa d’entrada:*

*Inputs (Entrades):* És la informació inicial que rep la neurona des de l’exterior. Les entrades poden ser tant binàries com contínues.

*Pesos (*)*:* Són coeficients que s’adapten dins de la xarxa i que determinen la intensitat de la senyal d’entrada registrada per la neurona, és a dir, són coeficients que depenent de la intensitat de l’imput per la resposta té un valor (pes) més elevat o menys.

Cada entrada té un pes que informa de la importància que té aquella entrada en la xarxa.

*- Capes ocultes:* La neurona consta de tres parts

*Regla de Propagació:* La seva funció principal és combinar les entrades i els pesos.

*Funció d’activació:* La neurona té un estat d’activació, això vol dir que pot estar activa o inactiva. La funció d’activació determina l’estat d’activació actual de la neurona.

La funció d’activació distorsiona el valor de sortida afegint deformacions no lineals. Hi ha diferents funcions d’activació

*Funció de sortida:* L’últim component de la neurona. Proporciona la sortida en funció de l’estat d’activació.

- Capa de sortida:

*Outputs (Sortides):* Ens mostren el resultat de la nostra xarxa neuronal.

Podem tenir més d’un output i pot ser tant continu com binari.

4.2.1 Neural Network Auto-Regressive (NNAR) model

Aquest model de xarxes neuronals es troba dins de les xarxes neuronals feedforward.

Les xarxes neuronals feedforward (directes) són les xarxes neuronals artificials més simples en que la informació només és mou cap a una direcció, des dels nodes d’entrada passant pels nodes ocults fins als nodes de sortida.

Les xarxes neuronals autoregressives, és un mètode molt utilitzat per la predicció de sèries temporals. Aquest mètode s’utilitza a l’R mitjançant la funció *nnetar()*.

Aquesta funció retorna un model del tipus on p indica els retards que s’utilitzen per la predicció, P indica el component estacional, k indica els nodes ocults i m indica la periodicitat (mensual, trimestral..). Quan les dades no tenen estacionalitat s’utilitza un model sense el paràmetre P.

Un model té i k neurones en la capa oculta.

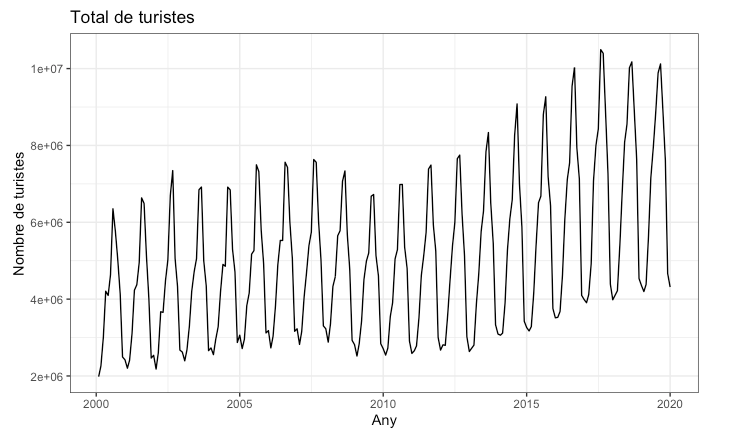
Un model és equivalent a un model però sense les resticcions dels paràmetres d’estacionalitat.

La funció nnetar() si no se li especifiquen els paràmetres s’encarrega automàticament. En aquest estudi, utilitzarem la funció nnetar() de manera automàtica i deixarem que l’R s’encarregui de fixar els paràmetres (ja que després utilitzarem un mètode manual). A més a més aquesta funció té un paràmetre *lambda* que s’encarrega de la transformació de Box-Cox. En el nostre cas utilitzarem utilitzarem el paràmetre com lambda=”auto”. D’aquesta manera l’R s’encarrega de trobar el millor paràmetre per així complir les assumpcions de normalitat.

De la mateixa manera que s’ha fet amb els mètodes manuals, cal posar a prova els nostres models. Utilitzarem la mateixa sèrie que hem utilitzat en els mètodes manuals.

En aquest cas, com ja s’ha realitzat anteriorment no farem ningun estudi previ de l’estacionalitat (ja que estem tractant les mateixes dades)

Recordem que les nostres dades tenen la següent forma:



Procedim a realitzar la modelització:

Primer de tot utilitzem com hem dit la funció nnetar() i obtenim els següents resultats:

Series: df

Model: NNAR(3,1,2)[12]

Call: nnetar(y = df, lambda = "auto")

Average of 20 networks, each of which is

a 4-2-1 network with 13 weights

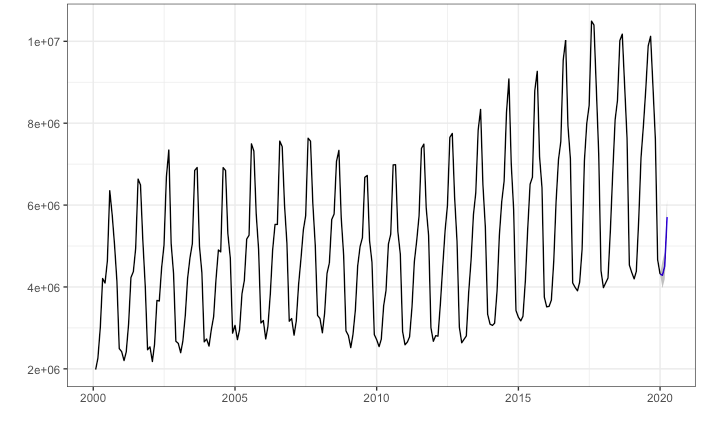
options were - linear output units

sigma^2 estimated as 0.09267

Hem obtingut un model NNAR(3,1,2) [12] això vol dir que el nostre model consta de 3 retards, 1 component estacional i 2 nodes ocults i amb una periodicitat mensual (12 mesos).

Ara, procedim a utilitzar el paquet forecast() per crear les prediccions i obtenim els següents resultats:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 4.284.967 | 4.507.412 | 5.707.169 |



Una vegada hem obtingut els valors reals podem calcular-los:

|  |  |  |
| --- | --- | --- |
|  | valor real | Predicció |
| Gener | 4.138.371 | 4.284.967 |
| ERROR |  | - 146.596 |

4.2.2 Long-Short Term Memory

5. Resultats i comparació

A continuació, es posarà en pràctica la teoria explicada anteriorment i els resultats significatius per l’anàlisi estaran presentats en taules.

Mètodes clàssics (ARIMA)

* Manual
* Automàtic (auto.arima)

Xarxes Neuronals

* Xarxes Neuronals Autoregressives
* Long-Short Term Memory

Tipologia

TOTAL

|  |  |  |  |
| --- | --- | --- | --- |
| pREDICCIÓ | Gener | Febrer | Març |
| Gener | 6.925.319  6.688.238  6.862.260 | 7.080.631  7.131.598  7.026.883 | 8.651.545  8.674.906  8.852.822 |
| Febrer |  | 7.062.832  7.182.899 | 8.633.435  8.761.825 |
| Març |  |  |  |

SARIMA(2,0,3)(2,1,1) [12]

SARIMA(3,0,3)(2,1,0) [12]

NNAR(5,1,4)[12]

|  |  |  |  |
| --- | --- | --- | --- |
| AIC | Gener | febrer | març |
| Gener | 6440.4  6423.3 | 6467.56  6450,72 |  |
| Febrer |  |  |  |
| Març |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | -53064  184017,3  9995 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

TURISTES

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 4.251.621  4.219.564  4.302.366 | 4.423.918  4.415.620  4.471.884 | 5.637.650  5.581.896  5.680.332 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(0,1,1)(1,1,1) [12]

SARIMA(0,1,1)(2,1,1) [12]

NNAR(3,1,2)[12]

|  |  |
| --- | --- |
| AIC | Gener |
| Gener | 6275.18  6267.89 |
| Febrer |  |
| Març |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | -113249.8  -81192.78  -163995 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

EXCURSIONISTES

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 2.650.772  2.665.996  2.690.152 | 2.669.307  2.656.996  2.677.616 | 3.074.446  3.064.110  3.004.172 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(1,1,1)(1,1,1) [12]

SARIMA(0,1,3)(0,1,1) [12]

NNAR(5,1,4)[12]

|  |  |
| --- | --- |
| AIC | Gener |
| Gener | 6096.55  6092.45 |
| Febrer |  |
| Març |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | 83112  67888  43732 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

Via d’accés

TOTAL

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 4.219.564  4.219.564  4.293.928 | 4.219.564  4.415.620  4.486.392 | 5.581.896  5.581.896  5.685.223 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(0,1,1)(2,1,1) [12]

SARIMA(0,1,1)(2,1,1) [12]

NNAR(3,1,2)[12]

|  |  |
| --- | --- |
| AIC | Gener |
| Gener | 6267.89  6267.89 |
| Febrer |  |
| Març |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | -81192.78  -81192.78  -155557 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

AEROPORT

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 3.590.574  3.578.298  3.647.053 | 3.681.015  3.677.240  3.707.546 | 4.722.456  4.707.977  4.693.519 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(1,1,1)(1,1,2)[12]

SARIMA(0,1,1)(0,1,1)[12]

NNAR(3,1,2)[12]

|  |  |
| --- | --- |
| AIC | Gener |
| Gener | 6146.69  6148.07 |
| Febrer |  |
| Març |  |

Aic model manual més baix???

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | -147754,8  -135478,6  -204234 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

CARRETERA

|  |  |  |  |
| --- | --- | --- | --- |
| predicció | Gener | Febrer | Març |
| Gener | 560.799,7  535.511,4  548.159,1 | 630.605.3  631.573,1  659.659,2 | 804.228,3  778.019,8  798.645,8 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(1,1,3)(0,1,1)[12]

SARIMA(1,0,2)(2,1,2)[12]

NNAR(3,1,2)[12]

|  |  |
| --- | --- |
| AIC | Gener |
| Gener | 5948,15  5966.2 |
| Febrer |  |
| Març |  |

Aic model manual més baix???

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | 17288,34  42576,63  29928.9 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

PORT

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 93.496,84  89.984,28  90.532,98 | 45.165,74  44.537,29  53.694,37 | 88.588,61  80.828,23  81.675,69 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(1,0,1)(0,1,0)[12]

SARIMA(2,0,2)(1,1,0)[12]

NNAR(3,1,2)[12]

|  |  |
| --- | --- |
| AIC | Gener |
| Gener | 5384,18  5380,05 |
| Febrer |  |
| Març |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | 10358.16  13870.72  13322,62 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

Comunitat Autònoma

TOTAL

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 4.243.462  4.219.564  4.284.967 | 4.442.793  4.415.620  4.5007.412 | 5.573.931  5.581.896  5.707.169 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(2,1,1)(2,1,1)[12]

SARIMA(0,1,1)(2,1,1)[12]

NNAR(3,1,2)[12]

|  |  |
| --- | --- |
| AIC | Gener |
| Gener | 6270,34  6267,89 |
| Febrer |  |
| Març |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | - 105.091  - 81.193  - 146.596 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

CATALUNYA

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 934.382,6  897.114  917.454,7 | 1.032.468,9  1.035.669  1.053.565 | 1.285.539,8  1.278.263  1.306.892,5 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(2,0,2)(2,1,0)[12]

SARIMA(1,0,1)(2,1,1)[12]

NNAR(3,1,2)[12]

|  |  |
| --- | --- |
| AIC | MODEL |
| Gener | 5850,3  5850,11 |
| Febrer |  |
| Març |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | -65214.64  -27946.04  -48286.7 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

MADRID

|  |  |  |  |
| --- | --- | --- | --- |
| Predicció | Gener | Febrer | Març |
| Gener | 617.143,1  610.375,1  619976,5 | 545.671,1  542.922,8  561732,1 | 683425,1  681.120,5  68995,3 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(0,1,1)(0,1,1)[12]

SARIMA(0,1,1)(1,1,2)[12]

NNAR(4,1,3)[12]

|  |  |
| --- | --- |
| AIC | MODEL |
| Gener | 5444,5  5435,32 |
| Febrer |  |
| Març |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | -5721,1  1046,9  -8554,5 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

COMUNITAT VALENCIANA

|  |  |  |  |
| --- | --- | --- | --- |
|  | Gener | Febrer | Març |
| Gener | 478.658,9  478.247,1  432.748,7 | 509.455,8  510.008,3  474.031,1 | 680.655,3  680.272,7  639.010,5 |
| Febrer |  |  |  |
| Març |  |  |  |

SARIMA(0,1,1)(1,1,1)[12]

SARIMA(0,1,1)(0,1,1)[12]

NNAR(3,1,2)[12]

|  |  |
| --- | --- |
| AIC | MODEL |
| Gener | 5443,74  5541,91 |
| Febrer |  |
| Març |  |

|  |  |  |  |
| --- | --- | --- | --- |
| ERROR | Gener | Febrer | Març |
| Gener | -44453.9  -44042.1  1456.3 |  |  |
| Febrer |  |  |  |
| Març |  |  |  |

6. Conclusions

Bibliografia

https://www.ine.es/dynt3/metadatos/es/RespuestaDatos.htm?oe=16028

<https://towardsdatascience.com/how-not-to-use-machine-learning-for-time-series-forecasting-avoiding-the-pitfalls-19f9d7adf424>

<https://www.cs.cmu.edu/afs/cs/academic/class/15782-f06/slides/timeseries.pdf>

<https://robjhyndman.com/hyndsight/nnetar-prediction-intervals/>

<https://www.cs.cmu.edu/afs/cs/academic/class/15782-f06/slides/timeseries.pdf>

<https://www.researchgate.net/publication/290106965_Neural_Networks_for_Time-Series_Forecasting>

<https://www.liip.ch/en/blog/time-series-prediction-a-short-comparison-of-best-practices?utm_name=iossmf>

<https://towardsdatascience.com/neural-networks-over-classical-models-in-time-series-5110a714e535>

<https://www.researchgate.net/publication/220668844_Feedforward_Neural_Nets_as_Models_for_Time_Series_Forecasting>

<https://www.researchgate.net/publication/324525859_Comparison_of_ARIMA_and_NNAR_Models_for_Forecasting_Water_Treatment_Plant's_Influent_Characteristics>

<https://otexts.com/fpp2/nnetar.html>

<https://pdfs.semanticscholar.org/93b7/9de5d49e26e933b5a731318390fe907d4957.pdf>

<https://www.statisticshowto.datasciencecentral.com/box-cox-transformation/>

<https://ca.wikipedia.org/wiki/Recessió_global_2008-2012>