

Creació d'una API que utilitzi Deep Learning

Juny 2018

Alumne: Ivan Martínez de la Torre
Director: Emilio Avilés Ávila
Ponent: Carme Quer Bosor
Especialitat d'enginyeria del software

Índex

1. Introducció	3
1.1. Context	3
1.2. Formulació del problema	3
1.3. Stakeholders	4
2. Estat de l'Art	5
2.1. Introducció	5
2.2. Neurones artificials	5
2.3. Com funcionen les xarxes neuronals artificials?	6
2.4. Com aprenen les xarxes neuronals artificials?	6
2.5. Tipus de xarxes neuronals	8
2.6. Deep Learning Frameworks	8
2.7. Conclusions	9
3. Abast	11
3.1. Fases i requisits del projecte.	11
3.2. Metodologia	12
3.3. Eines de seguiment	14
3.4. Mètode de validació	14
3.5. Aplicació de la metodologia	14
4. Planificació temporal	16
4.1. Delimitació temporal	16
4.2. Descripció de les tasques	16
4.2.1. Planificació del projecte	16
4.2.2. Anàlisi del software	17
4.2.3. Implementació del software	18
4.2.4. Programació web	19
4.2.4. Preparació final	19
4.3. Estimació temporal	19
4.4. Diagrama inicial de Gantt	21
4.5. Recursos utilitzats	22
4.5.1. Recursos humans	22
4.5.2. Recursos materials	24
4.5.3. Recursos software	26
4.6. Valoració d'alternatives i pla d'acció	28
5. Desenvolupament del projecte	30
5.1. Elecció del framework	30
5.2. Elecció del millors paràmetres	30

5.3. Selecció de les dades	33
5.3.1. Totes les dades	34
5.3.2. Dades individuals	34
5.3.3 Dades de les criptomonedes més antigues i importants	35
5.3.4 Conclusions	36
5.4. Creació de la web	36
5.4.1. Casos d'ús	36
5.4.2. Resultat de la web	37
6. Resultat de les prediccions	39
7. Autoavaluació del domini actual de la competència de sostenibilitat.	49
8. Pressupost	50
8.1. Estimació de costos	50
8.2. Control de gestió	52
8.3. Pressupost final	53
9. Informe de sostenibilitat	54
9.1. Estudi d'impacte ambiental	54
9.2. Estudi d'impacte econòmic	55
9.3. Estudi d'impacte social	56
9.4. Conclusions	56
10. Lleis i regulacions	58
11. Conclusions	59
11. Referències	60

1. Introducció

1.1. Context

Aquest és un Treball Final de Grau (TFG) que es realitza en la modalitat B amb un conveni de cooperació educativa a l'empresa SlashMobility [1]. Com a tal, hi ha un director de l'empresa i un ponent de la universitat. En el meu cas, el director del projecte a SlashMobility és Emilio Avilés Ávila, i la ponent és la professora del departament d'Enginyeria dels Serveis i Sistemes d'Informació Carme Quer.

SlashMobility va néixer al 2010 a Barcelona, capital mundial de la mobilitat i la innovació. És una empresa de solucions mòbils corporatives, la missió de la qual està centrada en accelerar la digitalització tecnològica de la societat mitjançant serveis mòbils 360° com: Formació, Disseny i Desenvolupament Software. Durant aquests anys s'ha convertit en una empresa de referència en el mercat del desenvolupament software d'aplicacions i en la formació en tecnologies mòbils. Actualment els +70 empleats, han desenvolupat +1000 solucions mòbils per a +150 empreses com: FCB, HP, Telefónica, Accenture, GFT, Santander, La Caixa, Desigual, Mango, Pronovias, Social Point, Letbonus, Habitaclia, Restalo, Gaesco, les quals han obtingut en el seu conjunt més de 10 milions d'euros de facturació, 100 milions de descàrregues, TOP 1 a les botigues globals principals, premis d'innovació, disseny i usabilitat. La facturació acumulada de Slashmobility supera els 6 milions d'euros, generant més de 40 milions de propostes i assessorant més de 350 companyies.

Fins ara Slashmobility s'ha centrat en solucions mòbils, però amb el recent desenvolupament de la intel·ligència artificial i de l'arribada de noves tecnologies com Blockchain, ha començat a investigar possibles noves vies per oferir serveis als seus clients.

1.2. Formulació del problema

Una de les línies d'investigació que interessa a Slashmobility és l'estudi de l'aplicació de les xarxes neuronals. Com a objectiu d'aquest projecte es buscava realitzar una prova de concepte del potencial de les xarxes neuronals en un repte difícil com és intentar predir l'evolució d'un mercat financer volàtil com és el mercat de les criptomonedes.

Per tal de proposar una solució a aquest repte, s'ha desenvolupat una primera versió d'una xarxa neuronal artificial (Deep Learning) [2]. Les xarxes neuronals són un dels paradigmes més interessants als que s'ha enfrontat el món de la programació en les últimes dècades. La lògica convencional en el món de la programació és dir-li a un ordinador que és el que ha de fer descomposant els problemes en petites tasques que aquest és capaç de realitzar. En contrast amb aquesta idea, en les xarxes neuronals artificials no són els programadors qui li diuen a l'ordinador què ha de fer, sinó que l'ordinador aprèn a partir de les dades existents trobant les seves pròpies solucions al problema.

El projecte que s'ha fet com a part del treball es basa en programar i entrenar amb les dades de l'històric de l'evolució del mercat de les criptomonedes una xarxa neuronal artificial que sigui capaç de preveure'n l'evolució futura. En aquesta versió no s'inclouen dades externes com podrien ser notícies relacionades amb les criptomonedes, ni dades d'altres mercats financers. Però en versions futures sí que s'afegiran per intentar realitzar prediccions més acurades.

1.3. Stakeholders

En aquesta secció es definiran els principals actors en el projecte i les persones o entitats interessades en els resultats del projecte. En definitiva, totes les persones o organitzacions que tenen algun tipus de relació amb el projecte.

SlashMobility

L'empresa ha estat representada per Emilio Avilés Ávila que s'ha encarregat de supervisar el projecte. Slashmobility s'ha beneficiat i es beneficiarà del projecte coneixent una nova tecnologia que podrà servir en un futur per obrir noves vies de serveis als seus clients.

Desenvolupador

Normalment, els desenvolupadors del projecte no s'inclouen de forma explícita com a stakeholders ja que generalment només s'encarreguen de realitzar el projecte. Però, en aquest cas, com que el projecte forma part d'un TFG, el desenvolupador també forma part dels stakeholders ja que també s'ha beneficiat i es beneficiarà perquè li servirà per aprendre una nova tecnologia i alhora per completar el grau i obtenir el títol universitari.

Usuaris

Tot i que el projecte està destinat a fer proves de concepte, qualsevol usuari podrà utilitzar la API resultant i predir pel seu compte el mercat de les criptomonedes.

Universitat

La Universitat està interessada en que els seus alumnes siguin capaços de realitzar bons projectes realitzant seguiment per part dels professors.

2. Estat de l'Art

Abans de començar el projecte el primer que es va fer és estudiar en quin estat de coneixement es troba actualment el món del Machine Learning, en concret, del Deep Learning. Per poder analitzar aquest coneixement, es van seguir les explicacions de Kirill Eremenko [3] i Hadelin de Ponteves [4], instructors d'un curs altament valorat a la plataforma de cursos online Udemy [5].

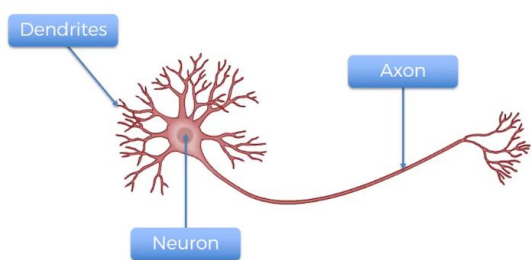
També es va estudiar l'estat de l'Art dels frameworks més comuns en el món del Deep Learning ja que com una de les tasques del projecte es va haver de seleccionar un d'aquests frameworks. Aquesta selecció es va fer agafant com a criteri la senzillesa del framework en línies de codi i la facilitat de trobar informació relacionada amb el seu ús.

2.1. Introducció

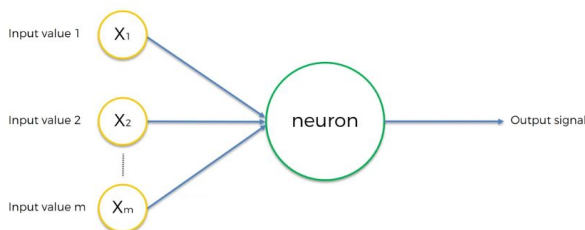
Primerament explicarem com varem afrontar l'estudi del Deep Learning. Es va començar estudiant què eren les neurones artificials, veient com aprenen i els algoritmes que usen. A continuació començarem parlant de les neurones (artificials). Després parlarem de com funcionen les xarxes neuronals artificials i de com aprenen. També parlarem dels diferents tipus de xarxes neuronals que hi trobem. I finalment, en la darrera secció comentarem els frameworks que s'utilitzen més comunament per resoldre aquest tipus de paradigmes.

2.2. Neurones artificials

Les neurones són el bloc més bàsic de les xarxes neuronals artificials. L'objectiu del Deep Learning és emular una neurona real i per fer-ho primer s'ha de conèixer que és una neurona. En la imatge següent (2.2.a), podem veure un esquema senzill d'una neurona:



2.2.a: Imatge d'una neurona real. (extret de [6])



2.2.b: Imatge d'una neurona artificial. (extret de [6])

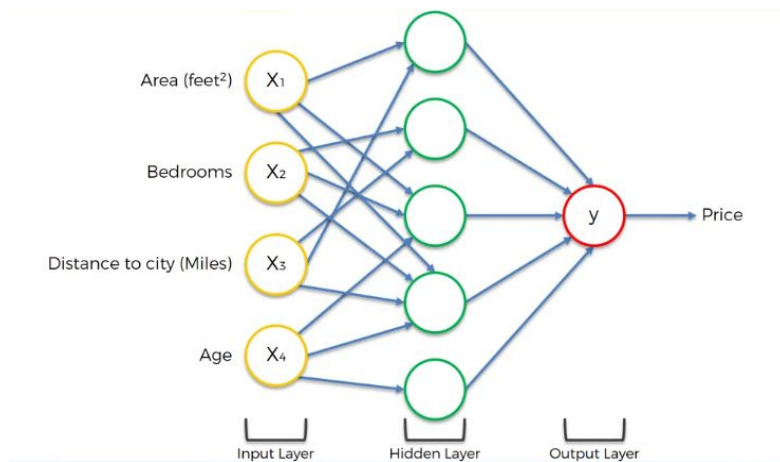
Les neurones tenen dues parts diferenciades, les dendrites i l'axó. Ambdós serveixen per comunicar-se amb altres neurones. Les dendrites són les encarregades d'obtenir informació provinent d'altres cèl·lules mentre que la funció de l'axó és transmetre la informació generada per la neurona a altres neurones. Ara que tenim una visió general de com funciona una neurona real, podem intentar definir com hauria d'actuar una neurona artificial. Per fer-ho ens ajudarem de la imatge anterior (2.2.b). Com podem observar, una neurona artificial hauria de ser capaç de rebre informació (inputs) i generar un resultat

(output) tal i com ho fan les neurones reals amb les seves dendrites i el seu axó. A efectes de programació, els inputs seran les nostres variables independents [6] a partir de les quals la nostra neurona generarà una variable dependent [7].

Cada neurona artificial té assignada una funció d'activació on suma els inputs que obté multiplicats per un pes específic. Cal esmentar que una neurona no té perquè tenir en compte tots els inputs que rep i pot utilitzar-ne només un subconjunt determinat. Els inputs han d'haver estat normalitzats i estandaritzats.

2.3. Com funcionen les xarxes neuronals artificials?

Una xarxa neuronal està formada per un conjunt de neurones. Fixant-nos en la imatge de més abaix (2.3), podem veure l'esquema més senzill d'una xarxa neuronal. Per una banda tornem a tenir els inputs i l'output i per altra banda tenim una capa de neurones (hidden layer a la imatge) on cada neurona treballa amb uns inputs determinats. Amb els outputs de les neurones es calcula el resultat.



2.3: Imatge d'una xarxa neuronal artificial. (extret de [6])

L'esquema exemple de la figura 2.3 fa referència a prediccions de preus de cases. Per una banda tenim els inputs (metres quadrats, nombre d'habitacions, distància al centre de la ciutat i anys d'antiguitat) i per l'altra banda tenim l'output (preu). Com hem comentat en la secció anterior, cada neurona treballa amb certs inputs, per exemple, la primera neurona treballa amb els metres quadrats de la casa i amb la distància envers el centre de la ciutat. Això vol dir que la xarxa neuronal ha trobat una correlació entre aquestes dues variables que influeix en el preu final de la casa. Aquest exemple es tracta d'una xarxa neuronal simple i a l'hora de la veritat no només hi ha una capa de neurones, sinó que n'hi han moltes, de manera que unes treballen amb els outputs de les altres fins que s'arriba a calcular l'output final.

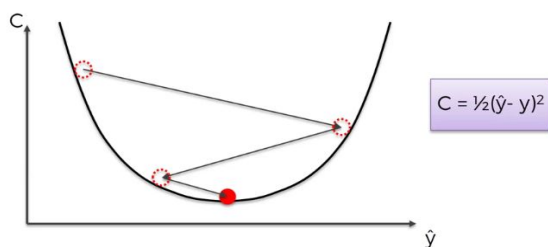
2.4. Com aprenen les xarxes neuronals artificials?

Quan un conjunt de inputs arriba a la xarxa neuronal, cada neurona de la primera capa calcula la seva funció (amb els inputs assignats a aquesta i els seus pesos). Els outputs de la primera capa passen a la segona i així fins que obtenim el output final. Un cop tenim

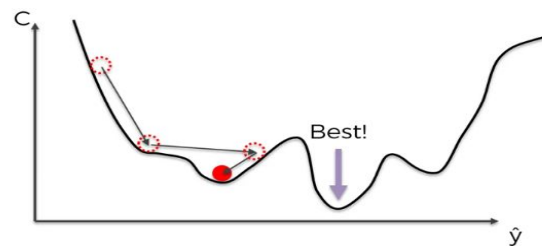
l'output final, es calcula la desviació (diferència entre el valor real i el valor predit per la neurona)^[8] del marge d'error de la predicció envers al resultat real i a partir d'aquesta desviació, la xarxa reajusta els pesos assignats als diferents inputs. Si es troba una nova manera de calcular possibles resultats, s'afegeix noves neurones a la xarxa. A continuació, explicarem més en detall com funcionen els algorismes de Stochastic Gradient Descent ^[9] i de Backpropagation ^[10].

Stochastic Gradient Descent

Abans de tot, parlarem del Normal Gradient Descent ^[11]. Aquest algorisme consisteix en agafar totes les entrades de dades i fer-les passar per la xarxa neuronal. Després sumar totes les desviacions generades per cada entrada, calcular el mínim fent zig zag com podem veure a la imatge **2.4.a** i actualitzar després els pesos dels inputs.



2.4.a: Batch gradient descent en funció convexa.
(extret de [6])



2.4.b: Batch gradient descent en funció no convexa.
(extret de [6])

Aquest algorisme té un problema, i ve representat per la imatge **2.4.b**, el normal gradient descent (o Batch gradient descent) només és fiable si la funció que segueix la desviació és convexa ja que sinó podem arribar a mínims parcials, i en el món real les funcions no acostumen a ser convexes. A partir d'aquesta observació, va sorgir l'anomenat Stochastic Gradient Descent, on en comptes de passar tota la base de dades alhora, es passa fila per fila i en ordres aleatoris. Després de passar cada fila, es calcula la desviació d'aquesta fila i s'actualitzen els pesos dels inputs. En el cas del Batch, sempre obtindríem el mateix resultat, però en el Stochastic sempre seran diferents. També, una de les diferències és que l'Stochastic és més ràpid que el Batch ja que no carrega tantes dades.

Per últim, comentar que existeix una solució intermèdia anomenada Minibatch Gradient Descent. Aquesta solució és una barreja de les dues explicades on el que es fa és formar paquets de files de 5, 10, 100... (a definir pel programador) i actualitzar els pesos dels inputs després de cada paquet.

Backpropagation

Fins ara hem parlat diversos cops sobre que els pesos dels inputs s'han de reajustar, és en aquest moment on entra en acció l'algorisme de Backpropagation. Backpropagation és un algorisme complex i avançat que ens permet actualitzar tots els pesos de manera simultània. Sense aquest algorisme, els pesos s'haurien d'ajustar manualment i independentment i seria difícil saber quin pes està provocant l'error en la desviació. Gràcies al backpropagation

podem saber quin o quins pesos estan afectant a l'error i ajustar-los tots alhora, no entrarem més en detall ja que les matemàtiques que hi han al darrera són bastant complexes.

2.5. Tipus de xarxes neuronals

La xarxa neuronal més bàsica consisteix en el què s'ha explicat fins ara. No obstant, existeixen dues evolucions més complexes per tal de resoldre problemes més específics que una xarxa neuronal artificial normal no pot. Aquestes dues evolucions són:

- Xarxes neuronals convolucionals: aquest tipus de xarxes utilitzen funcions convolucionals per poder reconèixer imatges, elaborar sistemes de recomanació i processat del llenguatge.
- Xarxes neuronals recurrents: aquest tipus de xarxes modifiquen les connexions bàsiques entre neurones habilitant cicles entre elles simulant, així, la memòria a curt plaç que tenen les xarxes neuronals humanes. Aquests tipus de xarxes s'utilitzen per intentar predir valors que depenen de cicles de temps.

2.6. Deep Learning Frameworks

En aquesta secció comentarem els diferents frameworks que podem trobar actualment per implementar solucions de Deep Learning. Aquests frameworks son TensorFlow [12], Theano [13], Keras [14], Torch [15] i Caffè [16]. A continuació anirem comentant-los un per un:

TensorFlow

TensorFlow és una llibreria *open source* per a computació numèrica que utilitza gràfics de flux de dades. Els diferents nodes del graf representen operacions matemàtiques, mentre que les arestes representen matrius multidimensionals de dades (tensors) que flueixen entre nodes. Aquesta arquitectura flexible ens permet desplegar computació en una o més unitats centrals de processament (CPU) o unitats de processament gràfic (GPU) en un ordinador, servidor o dispositiu mòbil sense reescriure codi.

TensorFlow va ser desenvolupat per investigadors i enginyers que treballaven per a Google's Machine Intelligence Research amb l'objectiu de portar el Machine Learning i el Deep Learning més enllà.

Theano

Theano és una llibreria *open source* de Python que ens permet definir, optimitzar i avaluar expressions matemàtiques que utilitzen matrius multidimensionals. Les característiques principals de Theano són:

- Bona integració amb NumPy.
- Ús transparent de GPU.
- Diferenciació simbòlica eficient.
- Possibilitat d'optimitzacions en velocitat i estabilitat.

- Generació dinàmica de codi en C.
- Capacitat de diagnòstic i detecció de molts tipus d'errors.

Keras

Keras és una llibreria *open source* sobre xarxes neuronals d'alt nivell, escrita en Python i capaç de funcionar sobre TensorFlow, CNTK o Theano. Les característiques principals de Keras són:

- Facilitat d'ús
- Totes les capes neuronals, funcions de desviacions, funcions d'activació, ... són mòduls i es poden combinar per crear models nous .
- És fàcil afegir mòduls nous.
- Treballa amb Python.

Torch

Torch és un framework *open source* que dona suport a una gran varietat d'algoritmes de Machine Learning i que prioritza l'ús de la GPU. És senzill d'utilitzar i eficient, gràcies a un llenguatge de scripting ràpid com és LuaJIT. Les característiques principals de Torch són:

- Una potent matriu N-dimensional.
- Moltes rutines per indexar, tallar, transposar, ...
- Interfície en C, a través de LuaJIT.
- Rutines d'àlgebra lineal.
- Xarxa neuronal i models basats en energia.
- Rutines d'optimització numèrica.
- Suport GPU ràpid i eficient.
- Ports a iOS i Android.

Caffe

Caffe és un framework *open source* per a deep learning que busca velocitat i modularitat. Les principals característiques de Caffe són:

- Arquitectura expressiva que permet canviar entre CPU i GPU només activant un flag.
- Codi extensible gràcies a la col·laboració de més de 1.000 desenvolupadors.
- Alta velocitat en processament d'imatges.
- Gran comunitat *open source* que hi contribueixen.

2.7. Conclusions

Segons el coneixement que es va poder adquirir durant tota la fase de investigació, es va seleccionar un tipus de xarxa neuronal i un framework que la implementés.

El tipus de xarxa neuronal més adequada per a l'objectiu d'aquest projecte és una xarxa neuronal recurrent. Com s'ha dit anteriorment, aquest tipus de xarxa neuronal és adequada per a intentar predir valors que depenen de cicles de temps, i aquest enfocament serà

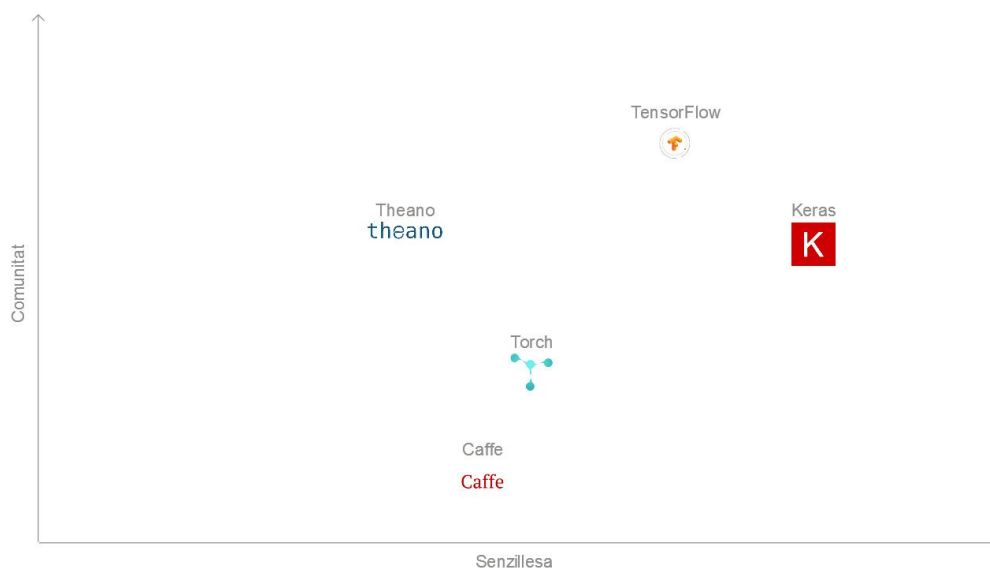
totalment adequat degut a que ens trobem en l'àmbit del mercat financer de criptomonedes i aquest depèn en gran mesura de cicles de temps.

Pel que fa al framework seleccionat, es va determinar segons proves de concepte entre els diferents frameworks esmentats. Aquesta decisió es va prendre durant la segona fase del projecte (veure capítol 3 i 4 d'aquesta memòria). Els criteris de selecció que varem usar van ser dos: continuïtat i senzillesa.

En quan a comunitat, es va avaluar la quantitat d'informació que vam poder reunir sobre el framework com la documentació oficial, el nombre de entrades a plataformes com StackOverflow i assumptes a Github.

En quan a senzillesa, es va valorar la quantitat de codi necessària per programar xarxes neuronals, també es van tenir en compte la claredat de conceptes en les documentacions oficials.

Un resum dels resultats de l'investigació es pot veure reflectit en la imatge 2.7. El resultat final va ser que el framework amb més comunitat al darrera és TensorFlow, i el més fàcil d'utilitzar és Keras. Després d'analitzar tots els frameworks ens vam decidir per Keras que es capaç de treballar per sobre de TensorFlow, i en definitiva ens val tota la informació trobada en ambdós frameworks.



2.7: Comparativa entre frameworks segons comunitat i senzillesa. [creació pròpia]

3. Abast

En aquest capítol s'expliquen les fases i requisits del projecte i com s'han cobert cadascuna i la metodologia de desenvolupament que s'ha seguit al llarg del projecte.

3.1. Fases i requisits del projecte.

En la taula següent definirem les diferents fases i requisits del projecte per tal d'especificar el seu abast. El projecte es va dividir inicialment en quatre fases. A continuació s'indiquen les fases planificades inicialment i es comenta el grau de compliment de la planificació per a cada fase del projecte.

Fase	Descripció inicial	Compliment
Investigació sobre Deep Learning	Una primera fase d'investigació sobre el Deep Learning per tal d'assolir coneixements sobre la matèria. Per fer-ho, és llegiran documents científics sobre aquesta àrea i es realitzarà un curs de Machine Learning i Deep Learning. Al finalitzar aquesta fase, hauríem de ser capaços de programar i entrenar una xarxa neuronal.	Aquesta fase s'ha completat correctament realitzant els cursos satisfactoriament i adquirint els coneixements necessaris per programar i entrenar una xarxa neuronal.
Anàlisi de les dades de les criptomonedes	En una segona fase del projecte, s'analitzarà la base de dades de l'històric del mercat de les criptomonedes i es programarà l'adaptació d'aquestes per tal de fer-les vàlides per entrenar la xarxa neuronal. També s'escollirà el framework a utilitzar mitjançant proves de concepte.	Les dades de les criptomonedes s'han normalitzat correctament. També s'han realitzat les proves dels diferents frameworks i s'ha escollit Keras amb TensorFlow. De totes les criptomonedes s'ha decidit treballar amb les deu més importants a nivell de volum de mercat.
Programació i entrenament de la xarxa neuronal	En una tercera fase del projecte, es programarà la xarxa neuronal i s'entrenarà amb les dades que hem obtingut a la fase anterior. També es crearà l'estructura del projecte amb el framework escollit en la fase anterior.	Durant aquesta fase s'han programat xarxes neuronals per les deu criptomonedes més importants.

Posada a prova de les xarxes neuronals	En la quarta fase es posarà a prova la xarxa neuronal fent que intenti predir els primers mesos de 2018 i comparant els resultats amb l'evolució real del mercat.	Durant aquesta fase s'han posat a prova totes les xarxes neuronals i s'han anat modificant fins obtenir els millors resultats possibles.
--	---	--

En la taula següent es defineixen, en la primera columna, els requisits que varem definir inicialment per a l'API que és objectiu del projecte. En la segona columna es pot trobar el grau de compliment d'aquests requisits al final del projecte i els resultats obtinguts.

Requisit	Compliment
L'API es realitzarà amb un framework gratuït.	100 % complert degut a que tant Keras com TensorFlow són frameworks open source.
R ² superior al 70% per a prediccions de valor únic.	<p>Els R² que s'han obtingut han estat:</p> <ul style="list-style-type: none"> ● Bitcoin: 83% ● Ethereum: 84% ● Ripple: 88% ● Bitcoin Cash: 90% ● Cardano: 89% ● NEO: 76% ● LiteCoin: 73% ● EOS: 72% ● NEM: 89% ● IOTA: 91% <p>Totes les criptomonedes han donat resultats superiors al 70%, per tant, aquest requisit queda 100% complert.</p>
Temps que triga en realitzar-se la predicció. Es considerarà complert si no supera el minut.	Els temps són inferiors als 30 segons. Per tant, requisit 100% complert.
L'API serà funcional i es podrà cridar per via http.	100% complert.

3.2. Metodologia

Avui dia les metodologies més comunes són les metodologies àgils. Aquest fet barrejat amb la limitació de temps que té un Treball Final de Grau (TFG) fa que ens decanem per metodologies àgils. SCRUM és una metodologia àgil i flexible per a gestió de desenvolupament de software, principal objectiu del qual és maximitzar el retorn de la

inversió per a l'empresa. Es basa en construir primer la funcionalitat de major valor per al client i en els principis d'inspecció contínua, adaptació, autogestió i innovació.

Amb la metodologia SCRUM el client s'entusiasma i es compromet amb el projecte donat que el veu créixer iteració a iteració. Així mateix, li permet en qualsevol moment realinear el software amb els objectius de negoci de la seva empresa, ja que pot introduir canvis funcionals o de prioritat a l'inici de cada nova iteració sense cap problema.

Els principals beneficis d'aquesta metodologia envers altres són:

- **Compliment d'expectatives:** El client estableix les seves expectatives indicant el valor que li aporta cada requisit, l'equip els estima i amb aquesta informació s'assigna la seva prioritat. Al final de la iteració es comproven els compliments dels requisits i es transmet la retroalimentació a l'equip.
- **Flexibilitat a canvis:** Alta capacitat de reacció davant els canvis de requeriments generats per necessitats del client o evolucions del mercat.
- **Reducció de l'arribada al mercat:** El client pot començar a utilitzar les funcionalitats més importants del projecte abans que es finalitzi completament.
- **Major qualitat del software:** La necessitat d'obtenir una versió funcional després de cada iteració, ajuda a l'obtenció d'un software de qualitat superior.
- **Major productivitat:** S'aconsegueix entre d'altres raons, gràcies a l'eliminació de la burocràcia i a la motivació de l'equip que proporciona el fet que siguin autònoms per a organitzar-se.
- **Maximitza el retorn de la inversió:** Producció de software únicament amb les prestacions que aporten major valor de negoci gràcies a la prioritització per retorn d'inversió.
- **Prediccions de temps:** Mitjançant aquesta metodologia es coneix la velocitat mitjana de l'equip per iteració, amb el què és possible estimar fàcilment per a quan es dispondrà d'una determinada funcionalitat.
- **Reducció de riscos:** El fet de dur a terme les funcionalitats de més valor en primer lloc i de conèixer la velocitat amb què l'equip avança en el projecte, permet aclarir riscos eficaçment de forma anticipada.

Tot i que el treball ha estat dut a terme per una sola persona, s'han pogut incorporar els conceptes àgils d'SCRUM, com són les iteracions setmanals o bisetmanals per tal d'anar adaptant el producte amb les necessitats dels stakeholders, que en aquest projecte han estat formats pel director del TFG per part de l'empresa.

S'aprofiten conceptes apresos a les assignatures d'enginyeria de requisits (ER) i de projecte d'enginyeria del software (PES), on primer es definien de forma concreta els objectius i requisits del projecte, s'acordaven amb els stakeholders (director del TFG en aquest projecte) i després es passava a la part de desenvolupament del producte de forma iterativa.

Sprints bisetmanals

De les metodologies àgils hem adoptat el concepte d'iteracions curtes, o sprints. Per al nostre projecte s'ha decidit que els sprints siguin bisetmanals i s'han culminat amb reunions amb el director del projecte al finalitzar cada iteració.

Reunions diàries

Cada dia s'han dedicat cinc minuts per a reunir-se amb el director del projecte i avaluar l'avanç del projecte. Aquest concepte és una adaptació de les daily meetings d'SCRUM. És una bona pràctica que creiem que ha fet que el projecte no es trobi amb tants problemes com podria.

3.3. Eines de seguiment

Per a dur a terme el projecte hem necessitat una sèrie d'eines per tal d'anar documentant i organitzant les diferents tasques del projecte. Per assolir aquesta idea, s'ha utilitzat Bitbucket com a repositori online de codi i Jira com a gestor de tasques i iteracions. El fet que ens decanem per aquestes dues eines és que es poden barrejar i a partir d'una tasca del Jira podem accedir al codi associat al Bitbucket. Addicionalment, també s'ha utilitzat Asana com a gestor de tasques no tècniques.

3.4. Mètode de validació

Per fer un seguiment general del projecte es va acordar amb el director que es realitzarien reunions diàries de 5 minuts per tal de no desviar el projecte. També es va acordar una reunió presencial més llarga cada dues setmanes per tal d'afrontar adequadament les diferents iteracions del projecte. A més a més, ambdós hem tingut sempre el correu electrònic disponible per possibles dubtes.

3.5. Aplicació de la metodologia

En aquesta secció valorarem l'aplicació de la metodologia descrita anteriorment, és a dir, compararem la teoria amb el que s'ha realitzat realment.

La metodologia que s'ha seguit ha estat SCRUM, tot i que no s'han seguit tots els punts descrits exactament:

- **Sprints bisetmanals:** tot i que la idea inicial era realitzar sprints de dues setmanes, hi han hagut problemes que en alguna ocasió han fet que un sprint s'allargués alguna setmana més. Com a contrapartida, en els sprints finals, els sprints han estat més curts perquè les tasques s'han realitzat amb més agilitat, i això ha fet que tanquem sprints abans.
- **Reunions diàries:** l'objectiu d'aquestes reunions era portar un seguiment fluït del desenvolupament del projecte. No obstant, no ha estat possible seguir aquesta pràctica degut a problemes externs al projecte. Però sempre que s'ha pogut realitzar la reunió, aquesta s'ha dut a terme, i normalment això ha passat un cop o dos per setmana.

- **Eines de seguiment:** pel que fa a les eines de seguiment, l'ús d'aquestes ha estat del 100% ja que s'han utilitzat tant Jira, com Bitbucket, com Asana. Tot i que Asana ha estat en menor mesura.
- **Mètode de validació:** en aquesta secció es definien les reunions bisetmanals i diàries ja comentades. Pel que fa als correus electrònics, sempre han estat disponibles i les respostes han estat molt ràpides.

4. Planificació temporal

4.1. Delimitació temporal

El projecte ha tingut una durada d'uns quatre mesos aproximadament. Aquest període es va iniciar amb el començament de l'assignatura GEP a finals de febrer i ha finalitzat al juny amb la defensa del TFG.

La defensa del TFG, com hem comentat, està situada temporalment a finals de juny, però la planificació va intentar que el projecte finalitzés a principis de juny per tal de preparar la defensa adequadament. No obstant, la planificació també ha tingut en compte els possibles imprevistos que podien sorgir i ha deixat marges de temps en les tasques per poder afrontar-los. Si aquestos imprevistos són nombrosos, el temps reservat per a la defensa del projecte es veuria reduït.

En la secció següent es defineixen les diferents tasques que formen part de la planificació.

4.2. Descripció de les tasques

En aquesta secció es defineixen les diferents tasques que apareixen al diagrama de Gantt de la secció 4.4. Abans però, cal recordar que aquest projecte es realitza seguint una metodologia àgil (SCRUM) que permet que les planificacions variïn segons com avanci el projecte i, per tant, la planificació original no s'ha complert al 100%, i s'indiquen els canvis que han anat sorgint.

4.2.1. Planificació del projecte

És una de les fases més importants del treball ja que aquí és on es van definir tots els aspectes del projecte en quan a què i com es realitzaria, des de l'abast i els seus objectius finals fins a la seva planificació detallada, el seu pressupost i un estat de l'art inicial.

Aquesta fase va estar formada bàsicament per les diferents entregues de GEP i un cop finalitzada seria molt més fàcil per al projecte fer-li el seguiment i implementar-lo. Aquesta fase va estar inicialment pensada per a que durés des del 25 de febrer fins al 9 d'abril i en global s'ha complert adequadament, tot i que algunes de les tasques sí que han patit variacions de temps respecte la planificació inicial.

Les diferents tasques que inclou aquesta fase són:

Abast

Aquesta tasca té una durada de 7 dies (des del 25/02/18 fins el 05/03/18). Durant aquests dies es defineixen el context, l'abast, i la metodologia a seguir en el projecte. L'objectiu final és que siguem capaços de saber què i com farem el projecte. Aquesta tasca es va realitzar seguint la planificació inicial.

Planificació temporal

Aquesta tasca té una durada de 7 dies (des del 05/03/18 fins el 13/03/18). Durant aquests dies es defineix la planificació del projecte, amb el seu diagrama de Gantt i les corresponents descripcions de les tasques que hi apareixen. També es defineixen els recursos que s'haurien d'utilitzar per dur a terme el projecte. Aquesta tasca va sortir de la planificació inicial i es va finalitzar amb un dia de retard, per tant, la tasca va durar 8 dies (des del 05/03/18 fins el 14/03/18).

Pressupost i sostenibilitat

Aquesta tasca té una durada de 5 dies (des del 13/03/18 fins el 19/03/18). Durant aquests dies es realitzen una anàlisi de la sostenibilitat i una estimació dels costos del projecte. Aquesta tasca va sortir de la planificació inicial i es va finalitzar amb 2 dies de retard, per tant, la tasca va durar 7 dies (des del 14/03/18 fins el 22/03/18).

Presentació preliminar

Aquesta tasca té una durada de 5 dies (des del 19/03/18 fins el 25/03/18). Durant aquests dies s'elabora una presentació oral de tres minuts enregistrada en vídeo. En aquesta gravació s'exposen els continguts de les fases anteriors. Aquesta tasca va sortir de la planificació inicial i es va finalitzar amb 1 dia de retard, per tant, la tasca va durar 6 dies (des del 22/03/18 fins el 29/03/18).

Entrega especialitat

Aquesta tasca té una durada de 6 dies (des del 25/03/18 fins el 01/04/18). Durant aquests dies es reflexiona i es justifica la tria de les competències tècniques del projecte i l'adequació del mateix a les característiques de l'especialitat. Aquesta tasca va modificar la planificació inicial i es va finalitzar amb 4 dies d'antelació, per tant, la tasca va durar 2 dies (des del 29/03/18 fins el 01/04/18).

Document final

Aquesta tasca té una durada de 7 dies (des del 01/04/18 fins el 09/04/18) i es realitzarà conjuntament amb l'anterior. L'objectiu és millorar els lliuraments realitzats fins aleshores amb els comentaris tant dels professors de GEP com de la ponent i generar un document únic final que inclogui tots els anteriors. Aquesta tasca es va realitzar seguint la planificació inicial.

4.2.2. Anàlisi del software

Durant aquesta fase es pren la majoria de decisions sobre el projecte. Decisions com quin tractament s'aplica a les dades de la base de dades de les criptomonedes o quin framework s'escolleix d'entre els definits a l'estat de l'art. També es fa un estudi extensiu sobre la matèria del TFG, en aquest cas, el Deep Learning.

Aquesta fase estava inicialment pensada per a que durés des del 9 d'abril fins el 3 de maig, però degut a l'avançament d'algunes tasques, la fase va acabar el 20 d'abril. Les diferents tasques que inclou són:

Investigació sobre Deep Learning

Aquesta tasca té una durada de 7 dies (des del 09/04/18 fins el 17/04/18). Durant aquests dies es realitzen cursos sobre Machine Learning i Deep Learning i s'entra en el detall dels algoritmes que apliquen les xarxes neuronals. També s'aprèn a programar i entrenar una xarxa neuronal. Aquesta tasca es va realitzar amb antelació i per tant no es van necessitar de dies per dur-la a terme. Aquesta investigació es va dur a terme durant la primera fase en hores no planificades per fer-ho.

Anàlisi de la base de dades

Aquesta tasca té una durada de 7 dies (des del 17/04/18 fins el 25/04/18). Durant aquests dies es realitza un estudi de la base de dades del mercat de les criptomonedes i es decideix quines dades són útils per a la xarxa neuronal i com es tracten per tal que siguin vàlides pels algoritmes. També es realitza el tractament de dades, normalitzant-les per afavorir l'execució dels algoritmes. Aquesta tasca es va realitzar en 2 dies (des del 09/04/18 fins el 12/04/18). no seguint així la planificació inicial.

Elecció del framework

Aquesta tasca té una durada de 7 dies (des del 25/04/18 fins el 03/05/18). Durant aquests dies es realitzen proves de concepte de tots els frameworks explicats a l'estat de l'art i s'escolleix aquell que millor s'adapta als interessos de la xarxa neuronal que hem de desenvolupar. Degut a la gran comunitat i la senzillesa del codi ens decantem per utilitzar Keras i TensorFlow. Aquesta tasca es va realitzar seguint la planificació inicial i va durar 7 dies (des del 12/04/18 fins el 20/04/18).

4.2.3. Implementació del software

Aquesta és la fase que inclou tota la càrrega pràctica del projecte. Aquí s'implementa, s'entrena i es posa a prova la xarxa neuronal. També és la fase que més inconvenients pot generar per la falta d'experiència en la matèria.

Aquesta fase estava inicialment pensada per a que durés des del 3 de maig fins el 27 del mateix mes, però l'entrenament de la xarxa neuronal no va ser tan acurat com esperàvem i es va haver d'iterar entre entrenament i posada a prova fins obtenir resultats satisfactoris, aquesta prolongació va fer que la fase durés des del 20/04/18 fins el 21/05/18). Les diferents tasques que inclou són:

Programació de la xarxa neuronal

Aquesta tasca té una durada de 7 dies (des del 03/05/18 fins el 11/05/18). Durant aquests dies es programa la xarxa neuronal seguint els coneixements apresos i el framework escollit en la fase anterior. Aquesta tasca es va realitzar seguint la planificació inicial i va durar 7 dies (des del 20/04/18 fins el 28/04/18).

Entrenament de la xarxa neuronal

Aquesta tasca té una durada de 7 dies (des del 11/05/18 fins el 19/05/18). Durant aquests dies s'entrena la xarxa neuronal seguint els coneixements apresos en la fase anterior i amb

les dades extretes de la base de dades del mercat de les criptomonedes. Aquesta tasca es va realitzar seguint la planificació inicial i va durar 7 dies (des del 28/04/18 fins el 06/05/18).

Posada a prova de la xarxa neuronal

Aquesta tasca té una durada de 7 dies (des del 19/05/18 fins el 27/05/18). Durant aquests dies es posa a prova la xarxa neuronal comparant els resultats que la xarxa neuronal prediu per al primer trimestre del 2018 amb els valors reals d'aquest. Aquesta tasca va durar més del planificat inicialment degut a que les prediccions no eren gaire acurades i es va haver d'iterar l'entrenament de la xarxa neuronal amb posades a prova fins aconseguir prediccions més acurades. Al final aquesta iteració va ocupar 14 dies (des del 06/05/18 fins el 21/05/18).

4.2.4. Programació web

Aquesta fase no estava prevista inicialment, però arribats al punt de finalització de la fase anterior amb un marge de 6 dies i restant només la memòria final, es va optar per programar una petita web per mostrar els resultats obtinguts per la xarxa neuronal en front de les dades reals. Aquesta fase va durar 7 dies (des del 21/05/18 fins el 28/05/18) i només inclou una tasca, la programació de la web.

4.2.4. Preparació final

Aquesta fase consisteix en unificar tota la documentació generada de forma ben estructurada i preparar la defensa del TFG. Es realitza un cop acabada la fase d'implementació, i inicialment anava a durar des del 27 de maig fins la data de la presentació final del projecte. Al final, aquesta fase va començar un dia tard però es va finalitzar correctament (des del 28/05/18 fins el 26/06/18).

4.3. Estimació temporal

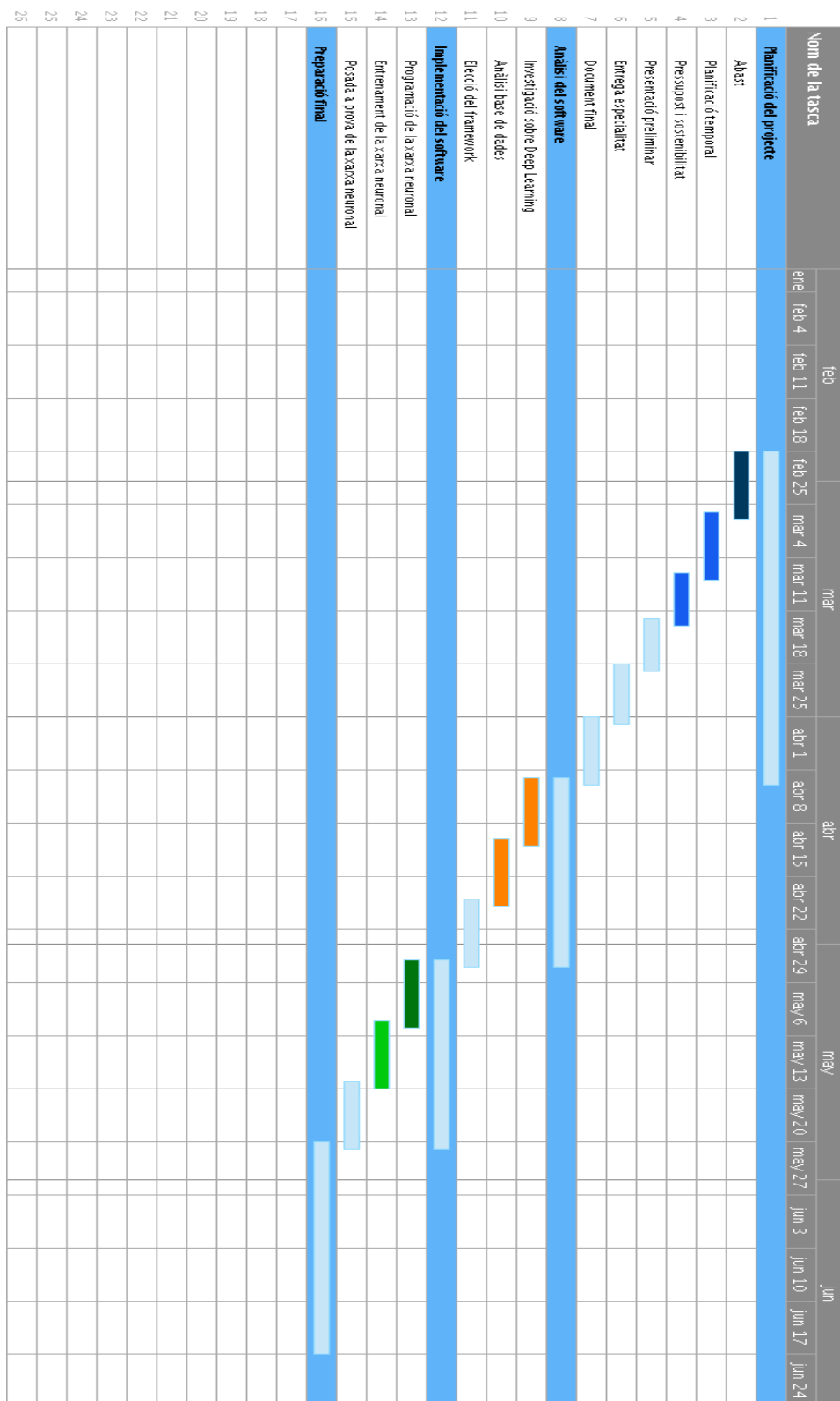
A continuació trobem un resum de les tasques del projecte amb les hores planificades inicialment i les hores que s'han utilitzat. Es pot observar que les hores de la fase de planificació del projecte i de preparació final tenen la meitat de dedicació que la resta ja que durant aquestes fases no s'utilitzen hores d'empresa per realitzar-les

Fase	Tasca	Hores planificació	Hores reals
Planificació del projecte	Abast	28	28
	Planificació temporal	28	32
	Pressupost i sostenibilitat	20	28
	Presentació preliminar	20	24
	Entrega especialitat	24	8

	Document final	28	28
Anàlisi del software	Investigació sobre Deep Learning	56	0
	Anàlisi de la base de dades	56	16
	Elecció del framework	56	56
Implementació del software	Programació de la xarxa neuronal	56	56
	Entrenament de la xarxa neuronal	56	56
	Posada a prova de la xarxa neuronal	56	112
Programació Web		0	56
Preparació final		28	28
Total		512	528

4.3.a. Taula comparativa d'estimació en hores per tasca i hores reals. [creació pròpia]

4.4. Diagrama inicial de Gantt



4.4.a. Diagrama de Gantt. [creació pròpia]

4.5. Recursos utilitzats

En aquest capítol es defineixen els recursos planificats i els que s'han utilitzat per realitzar el projecte, ja siguin humans, materials o software.

4.5.1. Recursos humans

Durant la realització del treball es necessiten diferents rols per dur a terme cadascuna de les tasques. Aquests rols són:

- **Cap de projecte (C):** Un cap de projecte, project manager o encarregat de projecte, és la persona que té la responsabilitat total del planejament i l'execució fins a l'acabament de qualsevol projecte.
- **Analista (A):** L'Analista és la persona que realitza les funcions d'anàlisi funcional i tècnica d'un projecte. Amb els seus coneixements es fixa la pila tecnològica i els requeriments funcionals.
- **Programador (Pg):** És un especialista en informàtica que és capaç d'elaborar sistemes informàtics, així com d'implementar-los, utilitzant un o més llenguatges de programació.
- **Provador (Pv):** És la persona que posa a prova un software amb l'objectiu de comprovar el seu funcionament i augmentar la seva qualitat.

A la taula següent observem el grau d'implicació dels diferents rols en el projecte segons el nombre d'hores que es van planificar que dedicarien a cada tasca:

Fase	Tasca	C (hores)	A (hores)	Pg (hores)	Pv (hores)	Hores
Planificació del projecte	Abast	20	8	-	-	28
	Planificació temporal	20	8	-	-	28
	Pressupost i sostenibilitat	10	10	-	-	20
	Presentació preliminar	10	10	-	-	20
	Entrega especialitat	14	10	-	-	24
	Document final	20	8	-	-	28
Anàlisi del	Investigació sobre	6	25	25	-	56

software	Deep Learning					
	Anàlisi de la base de dades	6	40	10	-	56
	Elecció del framework	6	30	20	-	56
Implementació del software	Programació de la xarxa neuronal	6	10	40	-	56
	Entrenament de la xarxa neuronal	6	10	40	-	56
	Posada a prova de la xarxa neuronal	6	5	5	40	56
Preparació final		14	14	-	-	28
Total		144	188	140	40	512

4.5.1.a. Taula estimació hores segons rols. [creació pròpia]

A la taula següent observem el grau d'implicació dels diferents rols en el projecte segons el nombre d'hores que s'han dedicat a cada tasca:

Fase	Tasca	C (hores)	A (hores)	Pg (hores)	Pv (hores)	Hores
Planificació del projecte	Abast	20	8	-	-	28
	Planificació temporal	22	10	-	-	32
	Pressupost i sostenibilitat	14	14	-	-	28
	Presentació preliminar	12	12	-	-	24
	Entrega especialitat	5	3	-	-	8
	Document final	20	8	-	-	28
Anàlisi del software	Investigació sobre Deep Learning	0	0	0	-	0
	Anàlisi de la base de dades	2	10	4	-	16
	Elecció del framework	6	30	20	-	56

Implementació del software	Programació de la xarxa neuronal	6	10	40	-	56
	Entrenament de la xarxa neuronal	6	10	40	-	56
	Posada a prova de la xarxa neuronal	12	5	45	50	112
Programació Web		6	5	45	-	56
Preparació final		14	14	-	-	28
Total		145	139	194	50	528

4.5.1.b. Taula d'hores reals segons rols. [creació pròpia]

4.5.2. Recursos materials

Durant la realització del treball s'han utilitzat diferents recursos materials per desenvolupar i implementar el projecte. Els recursos materials que s'han identificat són:

- **Ordinador portàtil (O):** És un dels recursos materials més importants i serveix per implementar el codi del projecte, per escriure la memòria del treball i realitzar les validacions.
- **Servidor (S):** És on s'allotja el projecte i roman en funcionament.
- **Espai d'oficina (EO):** És el lloc on es duu a terme moltes de les hores dedicades al projecte i està situat a les oficines de l'empresa.
- **Espai de l'habitatge (EH):** És el lloc de la residència personal on es duru a terme la resta d'hores que no formen part de les hores d'empresa dedicades al projecte.
- **Pantalla (P):** És l'element addicional de l'espai d'oficina on es pot connectar el portàtil per treballar més còmodament.

A la taula següent es defineix l'ús de cada recurs segons la planificació inicial i per tasca descrita prèviament:

Fase	Tasca	O	S	EO	EH	P
Planificació del projecte	Abast	Sí	No	No	Sí	No
	Planificació temporal	Sí	No	No	Sí	No
	Pressupost i sostenibilitat	Sí	No	No	Sí	No

	Presentació preliminar	Sí	No	No	Sí	No
	Entrega especialitat	Sí	No	No	Sí	No
	Document final	Sí	No	No	Sí	No
Anàlisi del software	Investigació sobre Deep Learning	Sí	No	Sí	Sí	Sí
	Anàlisi de la base de dades	Sí	No	Sí	Sí	Sí
	Elecció del framework	Sí	Sí	Sí	Sí	Sí
Implementació del software	Programació de la xarxa neuronal	Sí	Sí	Sí	Sí	Sí
	Entrenament de la xarxa neuronal	Sí	Sí	Sí	Sí	Sí
	Posada a prova de la xarxa neuronal	Sí	Sí	Sí	Sí	Sí
Preparació final		Sí	Sí	Sí	Sí	Sí

4.5.2.a. Taula estimació d'utilització de recursos materials segons tasques del Gantt. [creació pròpia]

A la taula següent es defineix l'ús real de cada recurs per tasca descrita prèviament:

Fase	Tasca	O	S	EO	EH	P
Planificació del projecte	Abast	Sí	No	No	Sí	No
	Planificació temporal	Sí	No	No	Sí	No
	Pressupost i sostenibilitat	Sí	No	No	Sí	No
	Presentació preliminar	Sí	No	No	Sí	No
	Entrega especialitat	Sí	No	No	Sí	No
	Document final	Sí	No	No	Sí	No
Anàlisi del software	Investigació sobre Deep Learning	Sí	No	Sí	Sí	Sí
	Anàlisi de la base de dades	Sí	No	Sí	Sí	Sí

	Elecció del framework	Sí	No	Sí	Sí	Sí
Implementació del software	Programació de la xarxa neuronal	Sí	No	Sí	Sí	Sí
	Entrenament de la xarxa neuronal	Sí	No	Sí	Sí	Sí
	Posada a prova de la xarxa neuronal	Sí	No	Sí	Sí	Sí
Programació Web		Sí	No	No	Sí	No
Preparació final		Sí	No	Sí	Sí	Sí

4.5.2.b. Taula real d'utilització de recursos materials segons tasques del Gantt. [creació pròpia]

4.5.3. Recursos software

Durant la realització del treball s'utilitzen diferents recursos software per desenvolupar i implementar el projecte. Els recursos software que s'han identificat són:

- **Bitbucket (B):** És un repositori online que permet emmagatzemar codi i portar control de versions de manera visual i senzilla.
- **Jira (J):** És un gestor d'iteracions i tasques que permet gestionar de manera eficient l'organització i temps de les fases del projecte.
- **Asana (A):** És un gestor de tasques més senzill i s'utilitza per portar un control de tasques no funcionals com programar reunions.
- **Google Drive (G):** És un software que permet emmagatzemar arxius i editar-ne d'alguns tipus. En aquest projecte s'utilitza per generar i guardar tots els documents necessaris per generar la documentació final.
- **Framework (F):** És el recurs que s'utilitza per programar i entrenar la xarxa neuronal.

A la taula següent es defineix l'ús de cada recurs segons la planificació inicial per cada tasca descrita prèviament:

Fase	Tasca	B	J	A	G	F
Planificació del projecte	Abast	No	No	No	Sí	No
	Planificació temporal	No	No	No	Sí	No

	Pressupost i sostenibilitat	No	No	No	Sí	No
	Presentació preliminar	No	No	No	Sí	No
	Entrega especialitat	No	No	No	Sí	No
	Document final	No	No	No	Sí	No
Anàlisi del software	Investigació sobre Deep Learning	No	No	Sí	Sí	No
	Anàlisi de la base de dades	No	Sí	Sí	Sí	No
	Elecció del framework	Sí	Sí	Sí	Sí	Sí
Implementació del software	Programació de la xarxa neuronal	Sí	Sí	Sí	Sí	Sí
	Entrenament de la xarxa neuronal	Sí	Sí	Sí	Sí	Sí
	Posada a prova de la xarxa neuronal	Sí	Sí	Sí	Sí	Sí
Preparació final		No	No	Sí	Sí	Sí

4.5.3.a. Taula planificació d'utilització de recursos software segons tasques del Gantt. [creació pròpia]

A la taula següent es defineix l'ús real de cada recurs per cada tasca descrita prèviament:

Fase	Tasca	B	J	A	G	F
Planificació del projecte	Abast	No	No	Sí	Sí	No
	Planificació temporal	No	No	Sí	Sí	No
	Pressupost i sostenibilitat	No	No	Sí	Sí	No
	Presentació preliminar	No	No	Sí	Sí	No
	Entrega especialitat	No	No	Sí	Sí	No
	Document final	No	No	Sí	Sí	No
Anàlisi del software	Investigació sobre Deep Learning	No	No	Sí	No	No

	Anàlisi de la base de dades	No	Sí	Sí	No	No
	Elecció del framework	Sí	Sí	Sí	No	Sí
Implementació del software	Programació de la xarxa neuronal	Sí	Sí	Sí	No	Sí
	Entrenament de la xarxa neuronal	Sí	Sí	Sí	No	Sí
	Posada a prova de la xarxa neuronal	Sí	Sí	Sí	Sí	Sí
Programació Web		Sí	Sí	Sí	No	No
Preparació final		No	No	Sí	Sí	No

4.5.3.b. Taula utilització real de recursos software segons tasques del Gantt. [creació pròpia]

4.6. Valoració d'alternatives i pla d'acció

En aquesta secció es mostren els riscos que van ser identificats durant la fase de planificació i s'explica si van aparèixer o no.

Escollir el framework incorrecte

Era possible que a priori penséssim que un dels frameworks era el millor i, per tant, decidíem que utilitzaríem aquest, però alhora de fer-l'ho servir en el nostre projecte ens adonéssim que no era vàlid per solucionar el problema. En aquest cas, hauríem d'haver canviat de framework i tornar a començar el projecte de nou.

Es va calcular que aquest risc tenia una probabilitat baixa de sorgir ja que per escollir el framework s'havien planificat força hores i proves de concepte. En cas que apareixes s'eliminarien la meitat d'hores de la fase de posada a prova de la xarxa neuronal per tal de corregir aquest risc.

Aquest risc era un dels més greus que podien sorgir i és per això que es van planificar tants esforços per definir el framework. Amb el projecte acabat podem afirmar que l'elecció del framework va ser correcta.

La xarxa neuronal no fa bones prediccions

Era possible que un cop entrenada la xarxa neuronal no fes les prediccions el més acurades possibles. En aquest cas, s'intentaria ajustar al màxim els paràmetres de la xarxa neuronal per tal de millorar les prediccions.

Es va calcular que aquest risc tenia una probabilitat mitjana de sorgir ja que els frameworks han estat utilitzats per molts desenvolupadors abans que nosaltres i de donar-se el cas significaria que l'entrenament no ha estat suficient. Per evitar aquest risc es van donar marge a les hores de posada a prova de la xarxa neuronal, així que la meitat d'hores d'aquesta fase servirien per seguir entrenant la xarxa. També es farien hores extres en l'habitatge personal.

Amb el projecte acabat podem afirmar que aquest risc va aparèixer i es va prendre la decisió d'allargar la fase de posada a prova de la xarxa neuronal gràcies a que es portaven bastants dies de marge de les fases anteriors. Tot i que el risc va aparèixer, se li va fer front adequadament.

Planificació errònia

Durant la realització del projecte ens podríem haver adonat que alguna de les tasques planificades inicialment podria no haver estat planificada acuradament i que necessites més temps per ser finalitzada, o en canvi, que en sobrès massa temps.

Es va calcular que aquest risc tenia una probabilitat baixa de sorgir ja que s'havien realitzat grans esforços per planificar les diferents fases del projecte. En el cas que s'hagués estimat a l'alça, no hi hauria problemes per acabar el treball i el temps sobrant es podria dedicar a perfeccionar la xarxa neuronal i/o a crear una petita web que fes de front-end de l'API creada. En canvi, en el cas que s'hagués estimat a la baixa s'ampliarien les dates de les fases, disminuint les hores dedicades a les fases de documentació i posada a prova de la xarxa neuronal. També es farien hores extres en l'habitatge personal.

Amb el projecte acabat podem afirmar que aquest risc ha aparegut en ambdós sentits, per una banda, hi van haver tasques que van durar menys del que s'havia planificat inicialment. Per altra banda, hi van haver tasques que van requerir més de temps per realitzar-se. A nivells globals, es va planificar a l'alça i per tant van sobrar dies de projecte, davant això es va decidir implementar una petita web per tal d'analitzar les prediccions de la base de dades amb les dades reals d'una forma més visual i entenedora.

Imprevistos

SlashMobility és una consultora, i com a tal té clients per als que desenvolupa i manté productes software. És possible que apareguessin problemes de gestió crítics que requerien la dedicació d'hores a altres projectes i els esforços dedicats al TFG es veiessin reduïts. En aquest cas, es substituirien les hores d'empresa per hores extres a l'habitatge personal. Es va calcular que aquest risc tenia una probabilitat baixa de sorgir ja que es porta un control eficaç de tots els projectes actuals de l'empresa.

Amb el projecte acabat podem afirmar que aquest risc no ha aparegut i no s'ha hagut de prendre cap tipus de mesura relacionada amb aquest risc.

5. Desenvolupament del projecte

En aquest capítol s'explicarà com ha transcorregut el projecte en quan decisions preses, resultats i evolucions al llarg de les fases del projecte.

5.1. Elecció del framework

Una de les primeres decisions important era escollir el framework que més encaix donés al problema proposat. Per fer-ho es van analitzar els diferents frameworks esmentats al capítol 2 seguint els criteris de selecció explicats. Els resultats, tal i com es mostren a la taula 5.1, denoten clarament que els frameworks més utilitzats són Tensorflow i Keras, amb una diferència important envers la resta pel que fa al nombre d'entrades i preguntes en la plataforma StackOverflow. El fet que hi hagi tanta comunitat resolent problemes sobre aquest frameworks facilita el desenvolupament per a nous desenvolupadors en l'àrea. Pel que fa al nombre de línies de codi aproximades per programar una xarxa neuronal no ha influït tant en l'elecció ja que una diferència d'unes 10 línies no es tradueix en gaire esforç extra.

Framework	Entrades StackOverflow	Línies aproximades
Tensorflow	29047	20
Theano	2364	30
Keras	8979	15
Torch	876	25
Caffe	2547	25

5.1 Taula d'anàlisi de frameworks. [creació pròpia]

Després d'aquest estudi es va arribar a la conclusió ja esmentada d'utilitzar Keras que permet treballar per sobre de Tensorflow.

5.2. Elecció del millors paràmetres

Per entrenar una xarxa neuronal s'han d'establir certs paràmetres. A continuació es defineixen els paràmetres configurables a partir de Keras i alguns de personalitzats:

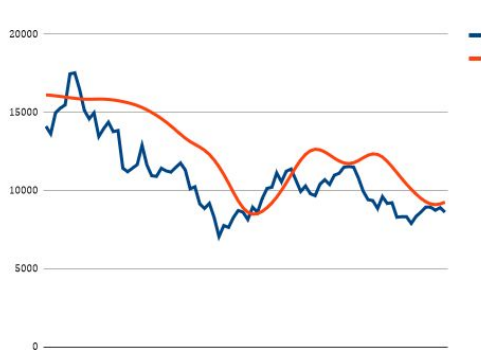
- **Nombre de dies de la xarxa recurrent:** Aquesta dada correspon al nombre de dies anteriors que s'agafen com a dades de entrada per a un dia concret.
- **Optimitzador:** Fa referència a l'algorisme de la xarxa neuronal per equilibrar els pesos de les diferents neurones.
- **Nombre de capes:** És el nombre de capes de neurones a través de les quals passen les dades abans de predir un resultat concret.
- **Nombre de neurones:** És el nombre de neurones que conté cada capa de la xarxa.

- **Epochs:** És el nombre de vegades que es passen les dades a la xarxa neuronal per tal d'entrenar-la.
- **Batch:** És el nombre d'entrades que entren a la xarxa paral·lelament.

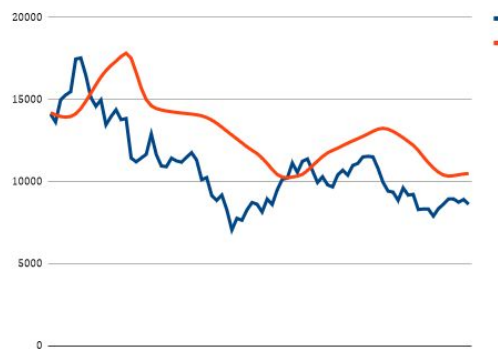
Per a escollir els valors dels diferents paràmetres es van realitzar els passos següents :

Primer es va decidir utilitzar només les dades del bitCoin per realitzar els ajustaments dels paràmetres. L'objectiu d'aquesta decisió era reduir el temps d'entrenament de la xarxa neuronal.

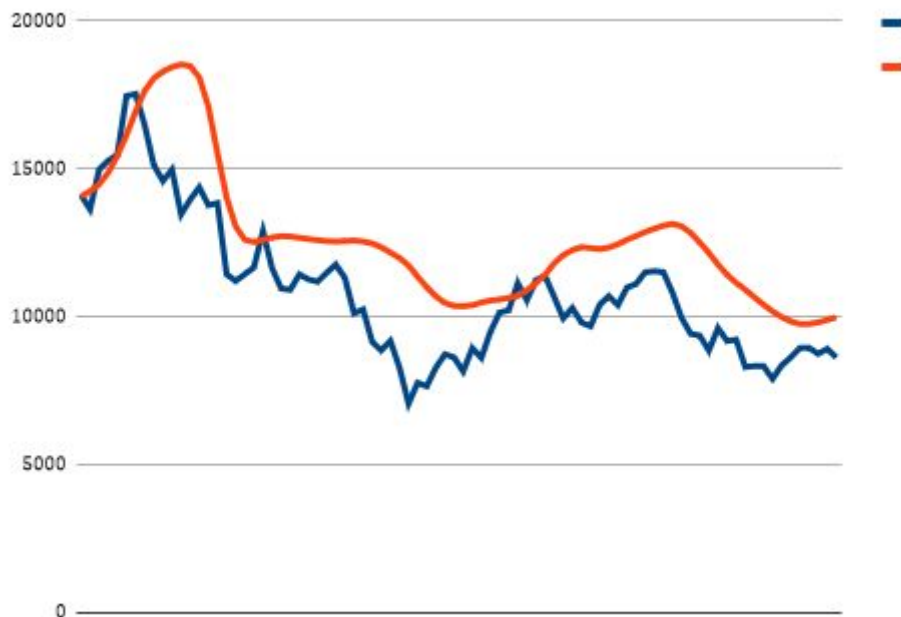
A continuació es van fixar tots els valors i es van realitzar proves d'entrenament modificant els paràmetres un per un. El primer valor a provar era el nombre de capes per el que passaran les dades. Es van realitzar proves amb 2, 4 i 8 capes. El resultat va ser:



5.2.a 8 capes (R² de 63%) [creació pròpia]



5.2.b. 2 capes (R² de 40%) [creació pròpia]



5.2.c 4 capes (R² de 72%) [creació pròpia]

Com es pot observar a les gràfiques **5.2.a**, **5.2.b** i **5.2.c** els resultats han estat de 63%, 40% i 72% de R² per a 8, 2 i 4 capes respectivament. Amb aquestes dades deduïm que el millor nombre de capes per aquest problema concret és **4**. A més a més, si ens fixem en les tendències de les línies de les prediccions (línies taronges), l'intent que més ha captat la tendència real (línies blaves) ha estat amb 4 capes de neurones.

Per a la resta de paràmetres es va utilitzar un avaluador que proporciona el framework Keras per triar la millor combinació de paràmetres. Aquesta funcionalitat permet definir els valors a provar i executa totes les combinacions. Keras retorna la combinació amb l'R² més elevat. Els valors que es van provar van ser:

Paràmetre	Valors
Nombre de dies	30, 60, 120
Optimitzador	Adam, RMSprop
Nombre de neurones	50, 100, 200
Epochs	50, 100, 200, 500
Batch	12, 25, 32, 56

5.2.d Taula de valors analitzats. [creació pròpia]

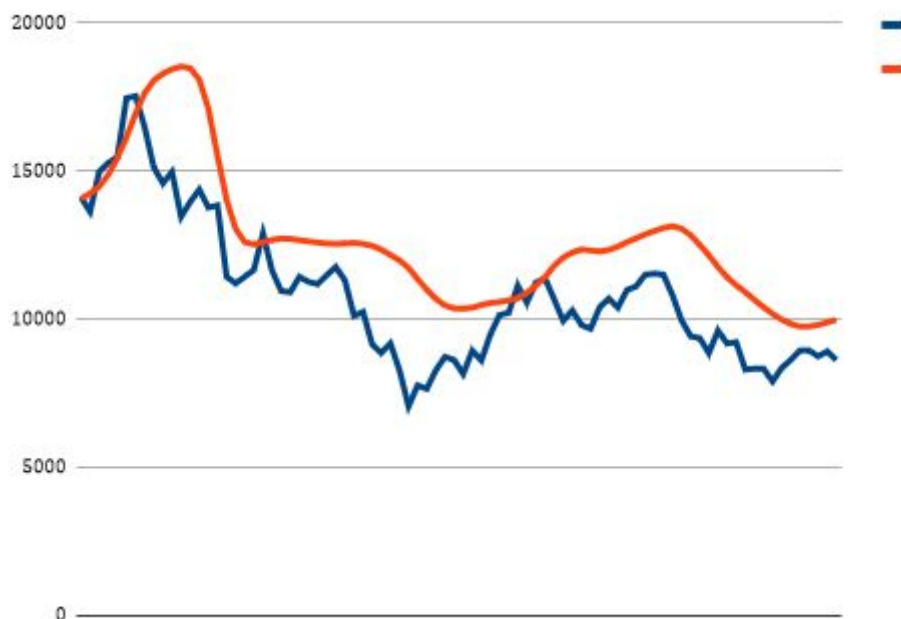
El resultat del avaluador més el paràmetre anterior va ser:

Paràmetre	Valor final
Nombre de dies	60
Optimitzador	Adam
Nombre de neurones	50
Epochs	100
Batch	32
Capes de neurones	4

5.2.e Taula de millors valors. [creació pròpia]

Un cop es van obtenir els paràmetres més òptims segons l'avaluador de Keras, es va predir la criptomoneda bitCoin amb l'objectiu de comprovar que l'R² obtingut era suficientment bo. A la gràfica **5.2.f** es pot observar aquesta predicció que va donar un R² de 72%. Donat que l'avaluador només es queda amb la millor combinació que troba no hem pogut conèixer els R² de la resta de combinacions. La possibilitat de provar totes les combinacions manualment es va descartar ja que el pròpi avaluador va trigar més de 3 dies sencers en realitzar-les totes. Tot i així els resultats són coherents i a continuació s'explica el perquè de cadascun:

- **Nombre de dies:** amb un nombre petit de dies la xarxa no té suficients variables per trobar correlacions entre elles. En canvi, massa dades fan que els pesos de les variables quedin tan repartits que no es creen tendències. És per això que l'avaluador va donar un resultat intermig en el nombre de dies.
- **Optimitzador:** es van provar els dos optimitzadors que proporciona Keras per a les xarxes neuronals recurrents. El resultat no ha estat analitzat ja que no s'ha entrat en detall en el funcionament d'ambdòs optimitzadors.
- **Nombre de neurones:** en aquest cas el resultat va ser millor per a capes amb menys neurones ja que fixen menys els càlculs que es realitzen i donen més marge per a possibles increments o decrements de tendències.
- **Epochs:** un altre cop un valor intermig va sortir com a paràmetre més òptim. En aquest cas és degut a que moltes iteracions de les dades donen menys marges per increments o decrements sobtats de tendències. En canvi, poques iteracions fan que les prediccions siguin poc acurades ja que no s'han analitzat suficient les dades.
- **Batch:** en aquest paràmetre creiem que qualsevol valor hagués donat un resultat similar ja que només afecta al nombre de dades que entren paral·lelament a la xarxa.



5.2.f Paràmetres òptims (R^2 de 72%) [creació pròpia]

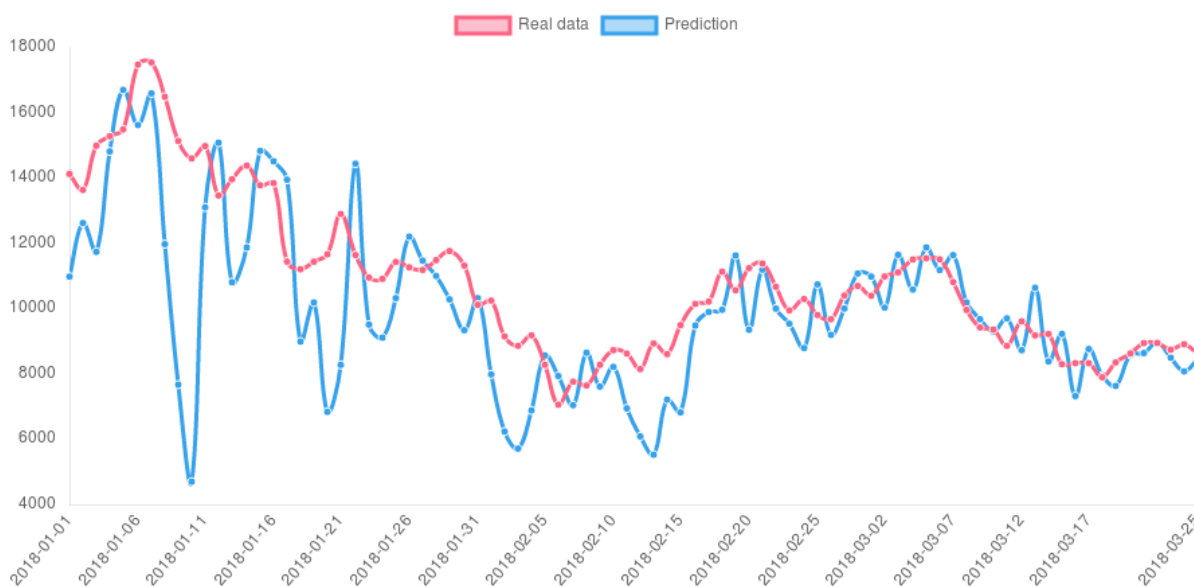
5.3. Selecció de les dades

Un cop el model de xarxa neuronal havia quedat definit es va procedir a seleccionar la millor combinació de dades. Per aquesta part, es va decidir provar a entrenar la xarxa neuronal de diferent maneres:

- Utilitzar totes les dades de totes les criptomonedes
- Entrenar una xarxa neuronal per cada criptomoneda amb les seves dades.
- Seleccionar un subconjunt de les criptomonedes en funció de l'antiguitat d'aquesta.

5.3.1. Totes les dades

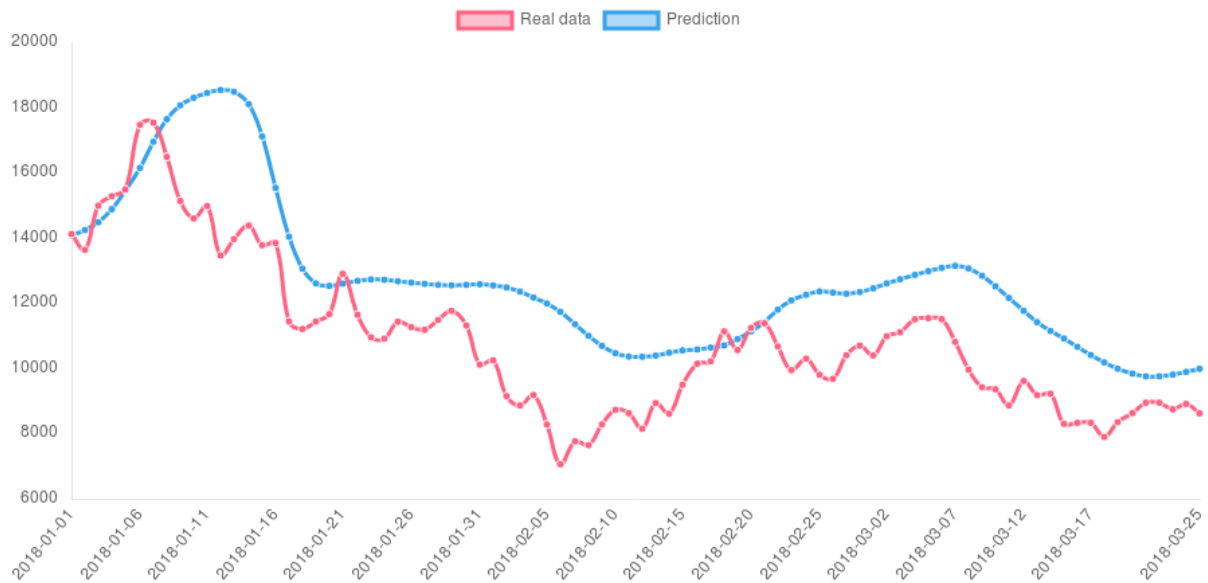
En aquesta prova es va entrenar la xarxa neuronal amb totes les dades de totes les criptomonedes disponibles. El resultat no va ser gaire bo ja que hi han moltes criptomonedes molt noves que no tenen cap recorregut i alteren les tendències. Com podem veure a la gràfica **5.3.1**, Les prediccions són bastant caòtiques generant massa alts i baixos que s'escapen massa de les dades reals. Un exemple important és la dada del 11 de gener on hi trobem una diferència de gairebé 10000 \$ entre la predicció i el valor real, aquest marge d'error és massa gran com per donar com a bona la predicció.



5.3.1 Predicció amb entrenament de totes les dades (R² de 47%) [creació pròpia]

5.3.2. Dades individuals

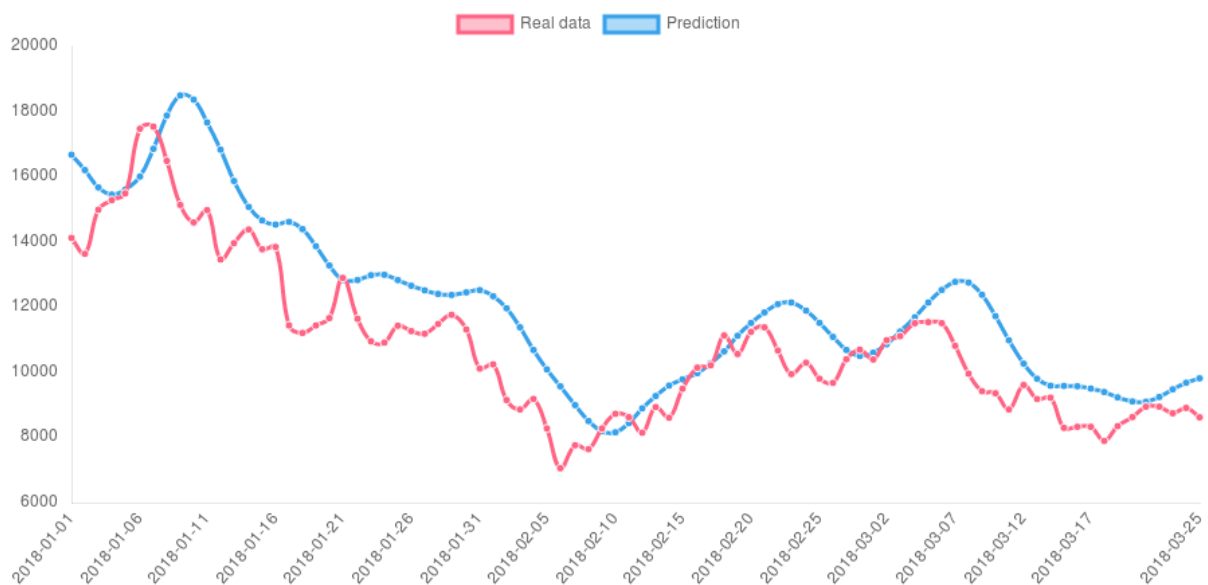
En aquesta prova es va entrenar una xarxa neuronal per cada criptomoneda. El resultat va ser bastant bo ja que cada moneda aprèn de les seves tendències pròpies. Com es pot veure a la gràfica **5.3.2**, la tendència que segueix la línia de la predicció s'aproxima bastant a la línia de les dades reals i el marge d'error es troba al voltant dels 2000 \$. Tot i que el marge d'error és alt la predicció de la tendència és acurada.



5.3.2 Predicció amb les dades de la pròpia criptomoneda (R^2 de 72% per al bitCoin) [creació pròpia]

5.3.3 Dades de les criptomonedes més antigues i importants

En aquesta prova s'utilitzen les dades de les criptomonedes més antigues i importants per entrenar la xarxa neuronal. Aquestes criptomonedes han estat bitCoin i Ethereum. El resultat ha estat molt bo ja que són les dues criptomonedes que més afecten al mercat financer de les criptomonedes degut que han ocupat entre les dues més del 80% del mercat. Com es pot observar a la gràfica **5.3.3**, la tendència de les prediccions és molt acurada, és a dir, detecta que el valor puja i baixa quan ha de fer-ho. El marge d'error volta els 1000 \$, en comparació amb les proves anteriors aquest és el millor marge d'error.



5.3.3 Predicció amb les dades de bitCoin i Ethereum (R^2 de 83%) [creació pròpia]

5.3.4 Conclusions

Un cop realitzades totes les proves anteriors es decideix a utilitzar només les dades de bitCoin i Ethereum per entrenar la xarxa neuronal. Aquesta decisió es fixa per la resta del projecte ja que no es creu que es puguin millorar més les prediccions amb les dades en disposició.

5.4. Creació de la web

Degut a que el projecte base es finalitza amb un petit marge de temps es decideix implementar una web que mostri els resultats obtinguts per la xarxa neuronal. Aquesta web es tracta d'una web simple sense disseny que té com a objectiu mostrar informació.

La web es programa en Angular 5 més Bootstrap 4 degut al coneixement que es té sobre aquests frameworks. Aquesta web conté:

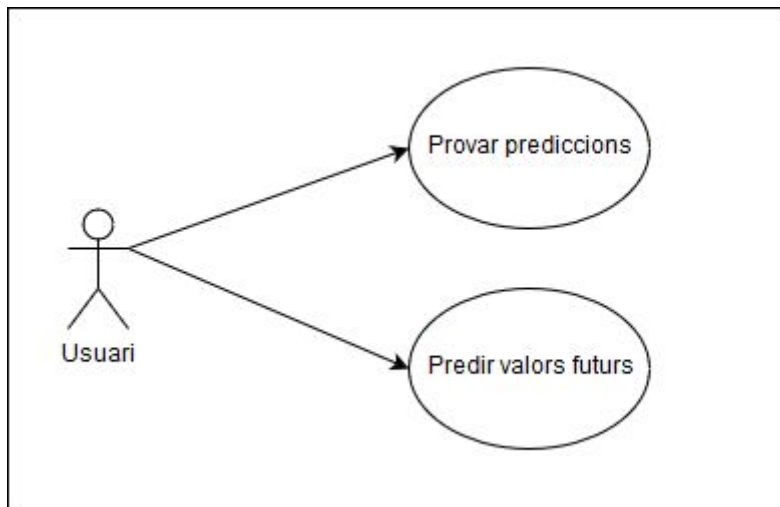
- Un selector de les deu criptomonedes més importants actualment.
- Una secció per posar a provar la xarxa neuronal envers les dades reals mitjançant prediccions a valors únics.
- Una secció per poder predir qualsevol data futura.

5.4.1. Casos d'ús

A continuació es defineixen els casos d'ús de les funcionalitats que podran utilitzar els usuaris:

- Prova de prediccions:
 - **Nom:** Prova de prediccions
 - **Responsabilitat:** Un usuari podrà consultar les prediccions a valor únic que realitza la xarxa neuronal.
 - **Precondicions:** L'usuari indica una criptomoneda de les top 10 del mercat criptofinancer.
 - **Postcondicions:** Se li mostrarà a l'usuari un gràfic comparatiu entre els valors reals i els valors predits per la xarxa. També se li mostrarà l' R^2 obtingut.
 - **Excepcions:** La criptomoneda indicada no existeix o no consta en les dades amb que s'ha entrenat la xarxa neuronal.
 - **Funcionament:** La funcionalitat crida al model de Keras i realitza les prediccions al moment. Per provar la xarxa s'agafen els últims 60 dies més els valors reals fins al 25 de març de 2018.
- Prediccions futures:
 - **Nom:** Prediccions futures
 - **Responsabilitat:** Un usuari podrà consultar les prediccions de la xarxa neuronal fins a una data concreta.

- **Precondicions:** L'usuari indica una criptomoneda de les top 10 del mercat criptofinancer.
- **Postcondicions:** Se li mostrarà a l'usuari un gràfic amb els valors predits per la xarxa neuronal.
- **Excepcions:** La criptomoneda indicada no existeix o no consta en les dades amb que s'ha entrenat la xarxa neuronal.
- **Funcionament:** La funcionalitat crida al model de Keras i realitza les prediccions al moment. S'agafen els últims valors dels 60 dies anteriors per predir el primer dia. Després es comencen a utilitzar els valors resultats de les prediccions per realitzar les següents fins arribar al dia indicat per l'usuari.

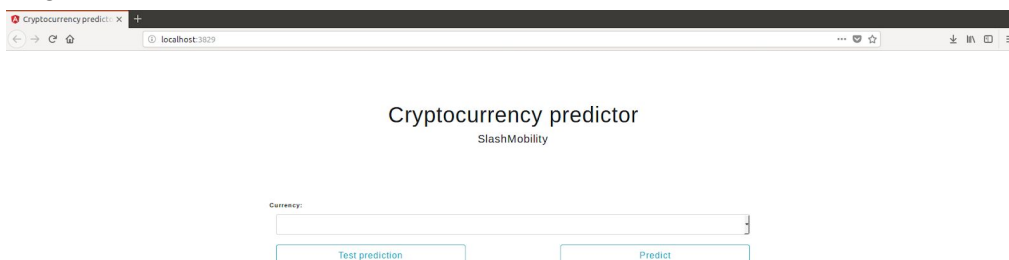


5.4.1 Diagrama de casos d'ús [Creació pròpia]

5.4.2. Resultat de la web

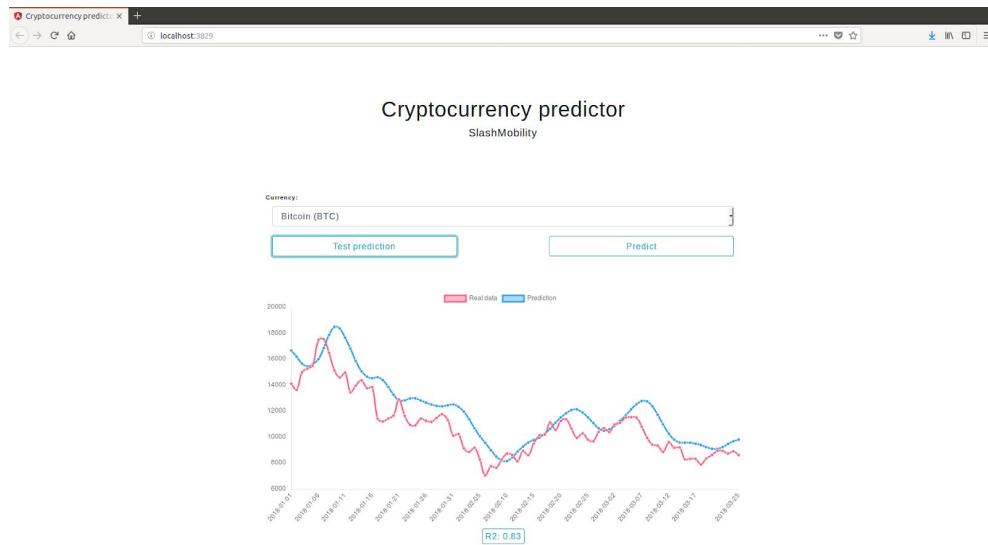
El resultat de la programació de la web es mostra en format de captures de pantalla de totes les vistes:

- Pàgina principal:



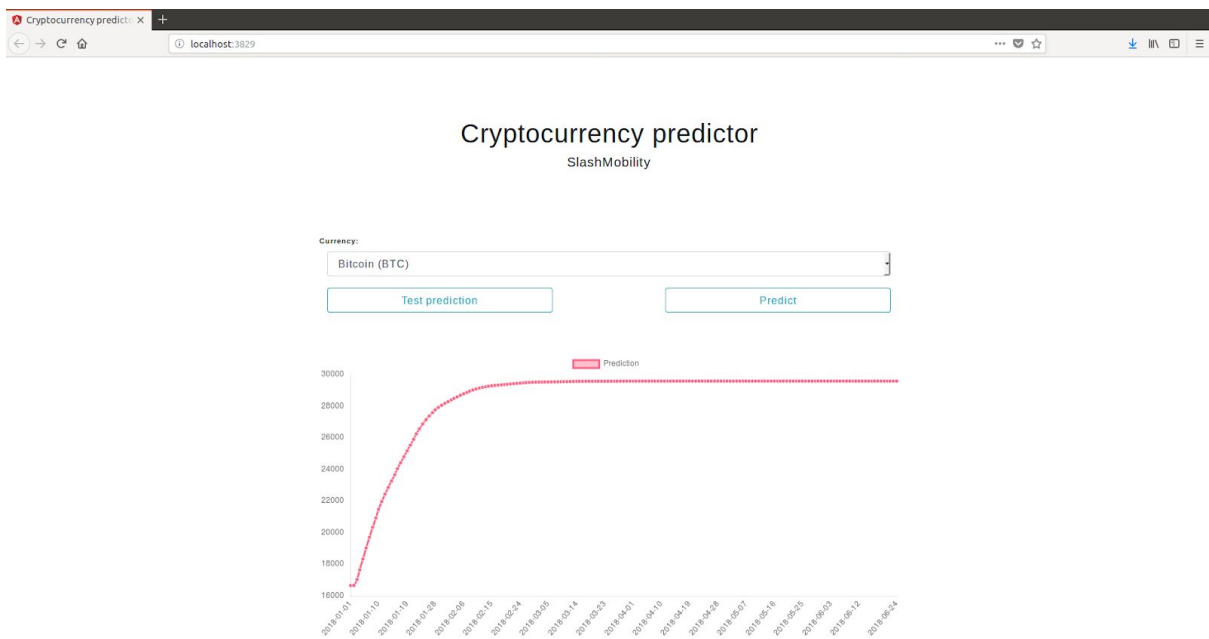
5.4.2.a Pàgina principal de la web [creació pròpia]

- Secció de prediccions de valors únics:



5.4.2.b Secció de prediccions de valor únic. [creació pròpia]

- Secció de prediccions futures:

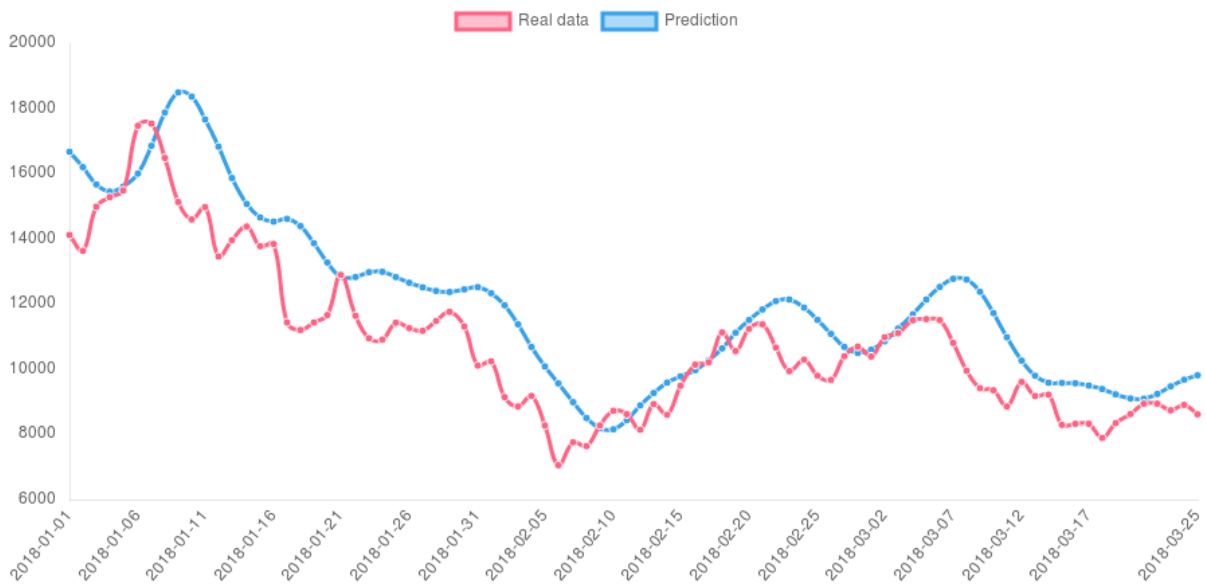


5.4.2.c Secció de prediccions futures. [creació pròpia]

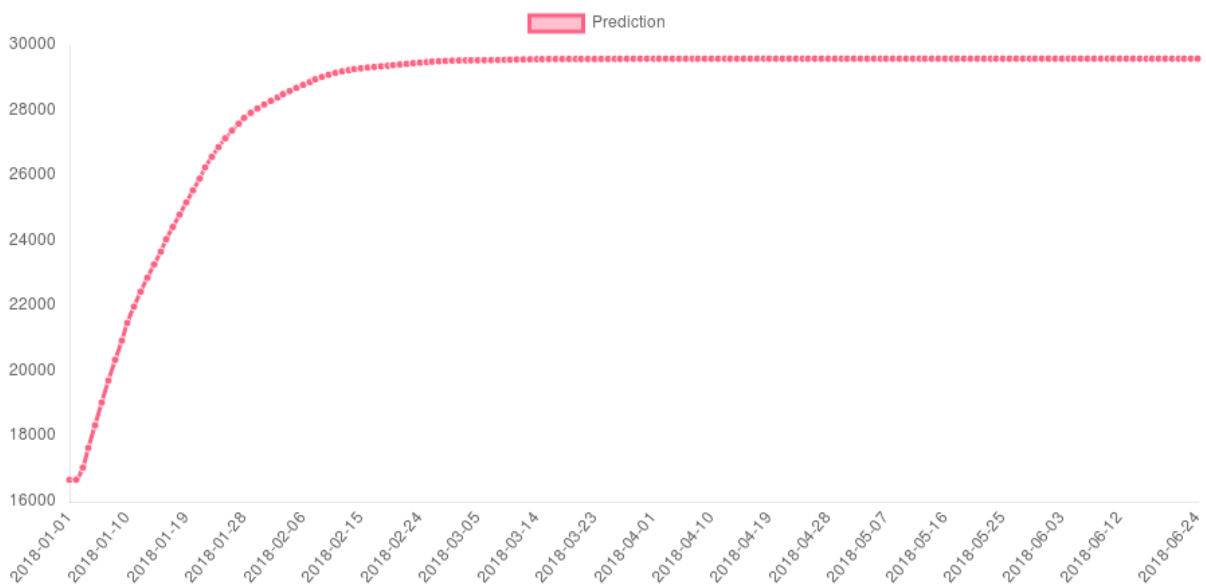
6. Resultat de les prediccions

En aquest capítol es mostren els resultats de totes les prediccions de valors únics i la comparativa amb les prediccions futures.

bitCoin



6.a Prediccions a valor únic bitCoin (R^2 de 83%) [creació pròpia]

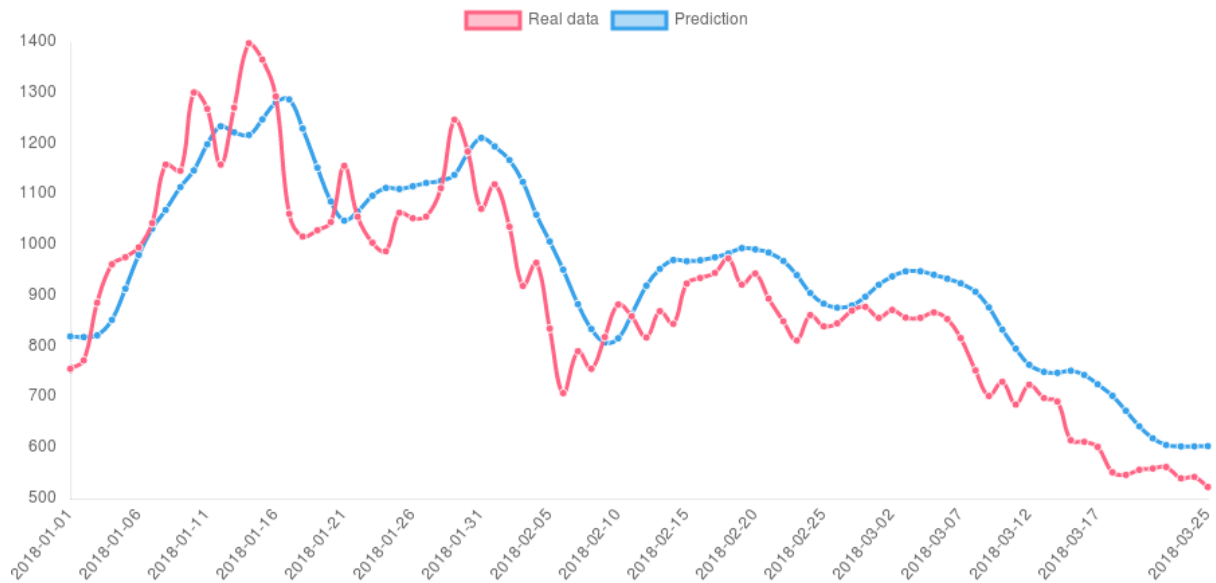


6.b Prediccions futures bitCoin (fins el 25 de juny de 2018) [creació pròpia]

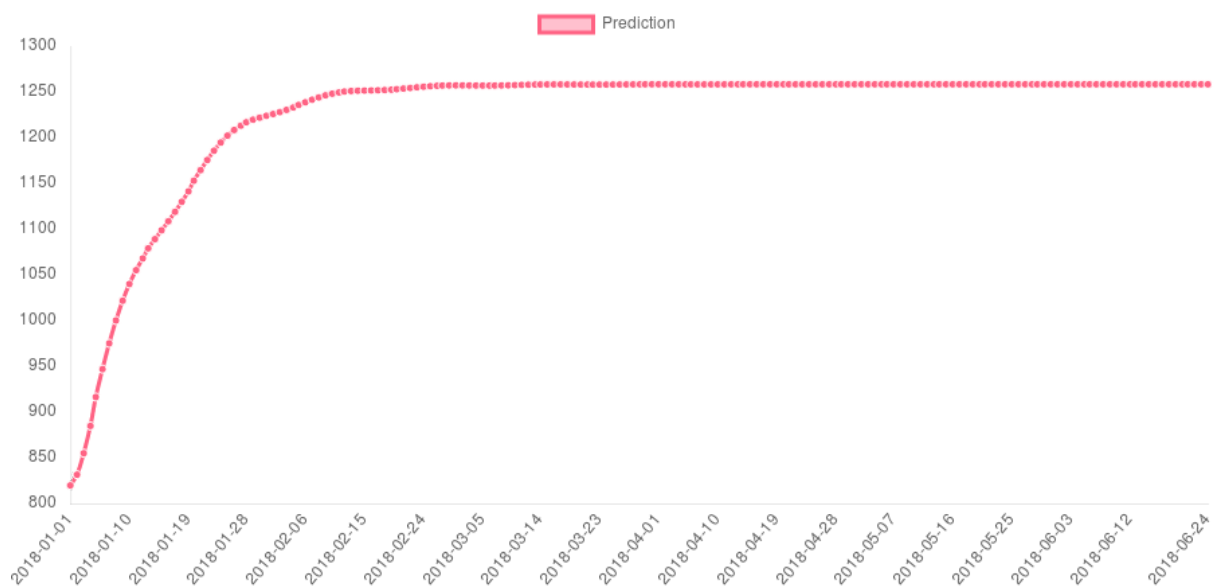
En la gràfica 6.a podem observar que el marge d'error entre els valors reals i els valors predits no supera els 1000 \$ per a valors que es mouen entre els 8000 \$ i els 18000 \$. A la

gràfica 6.b es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 30000 \$. Per últim es pot veure que l'R^2 d'aquesta predicció és del 83%.

Ethereum



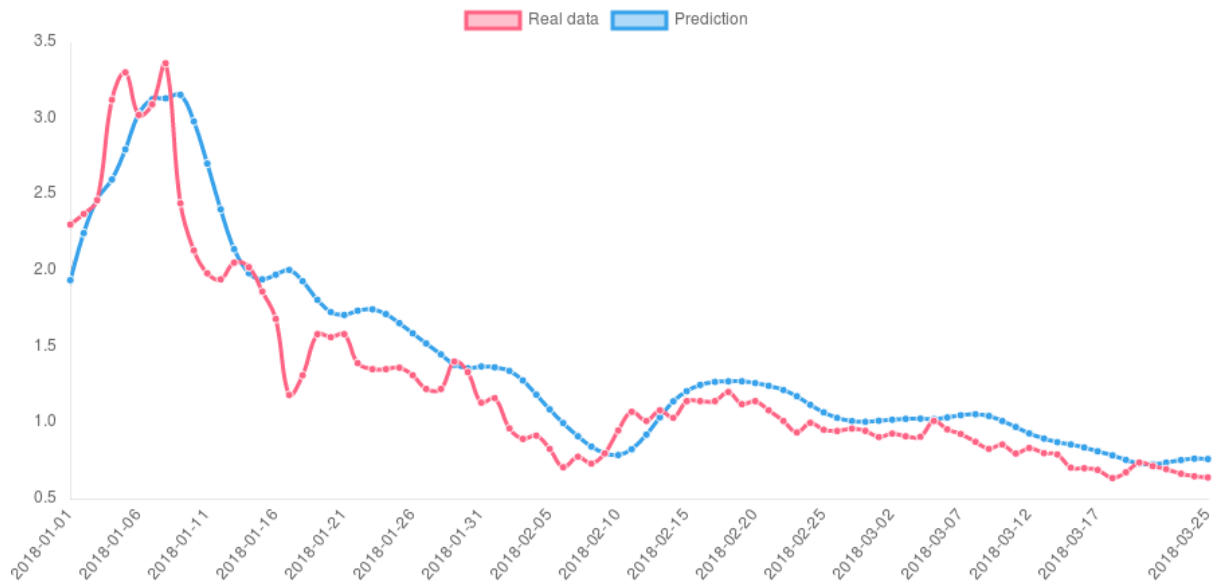
6.c Prediccions a valor únic Ethereum (R^2 de 84%) [creació pròpia]



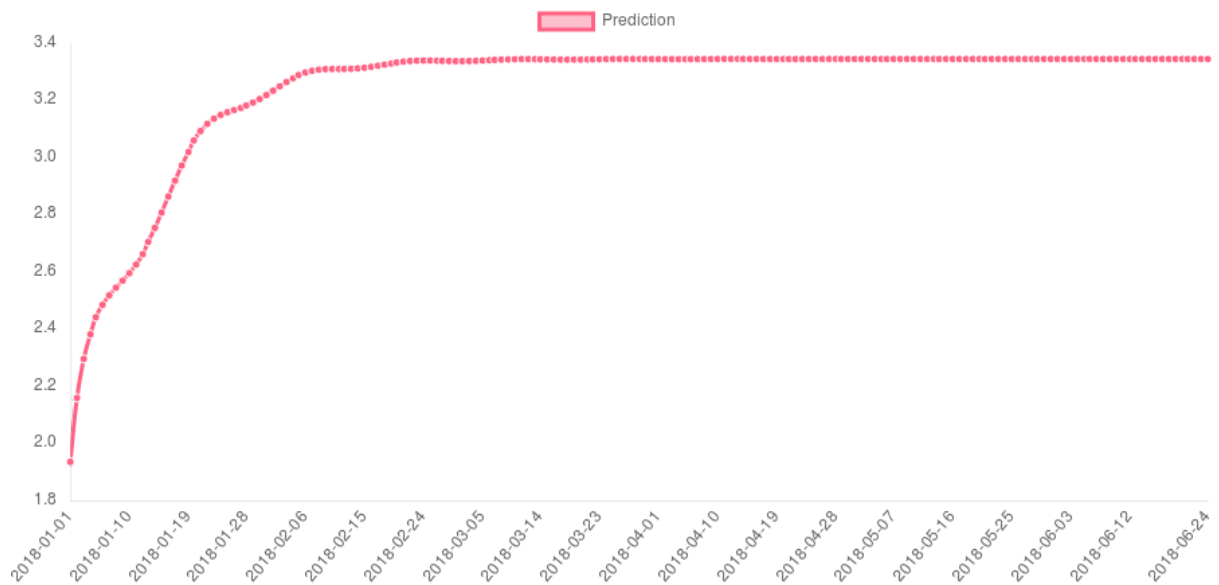
6.d Prediccions futures Ethereum (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica 6.c podem observar que el marge d'error entre els valors reals i els valors predits no supera els 100 \$ per a valors que es mouen entre els 500 \$ i els 1400 \$. A la gràfica 6.d es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 1250 \$. Per últim es pot veure que l'R^2 d'aquesta predicció és del 84%.

Ripple



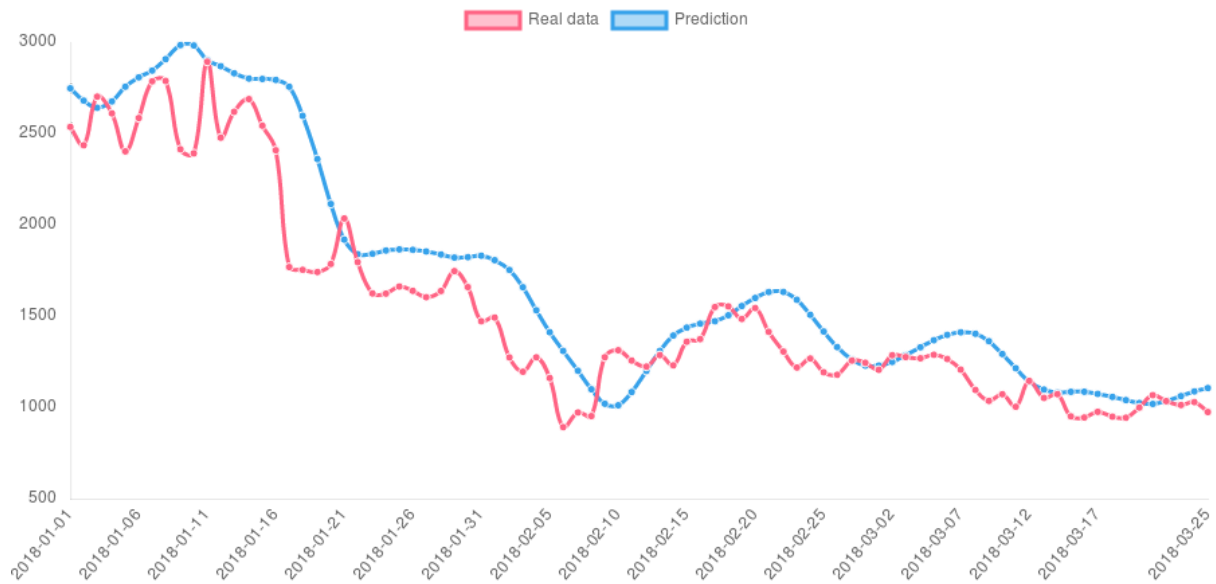
6.e Prediccions a valor únic Ripple (R^2 de 88%) [creació pròpia]



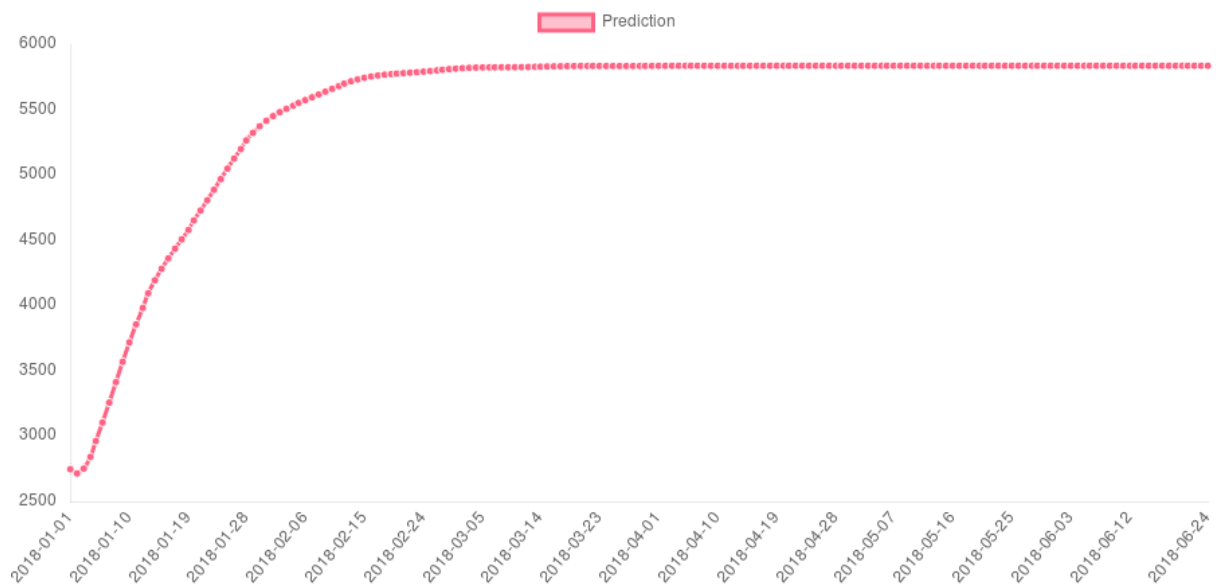
6.f Prediccions futures Ripple (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica 6.e podem observar que el marge d'error entre els valors reals i els valors predits no supera els 0.5 \$ per a valors que es mouen entre els 0.5 \$ i els 3.5 \$. A la gràfica 6.f es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 3.3 \$. Per últim es pot veure que l' R^2 d'aquesta predicció és del 88%.

bitCoin Cash



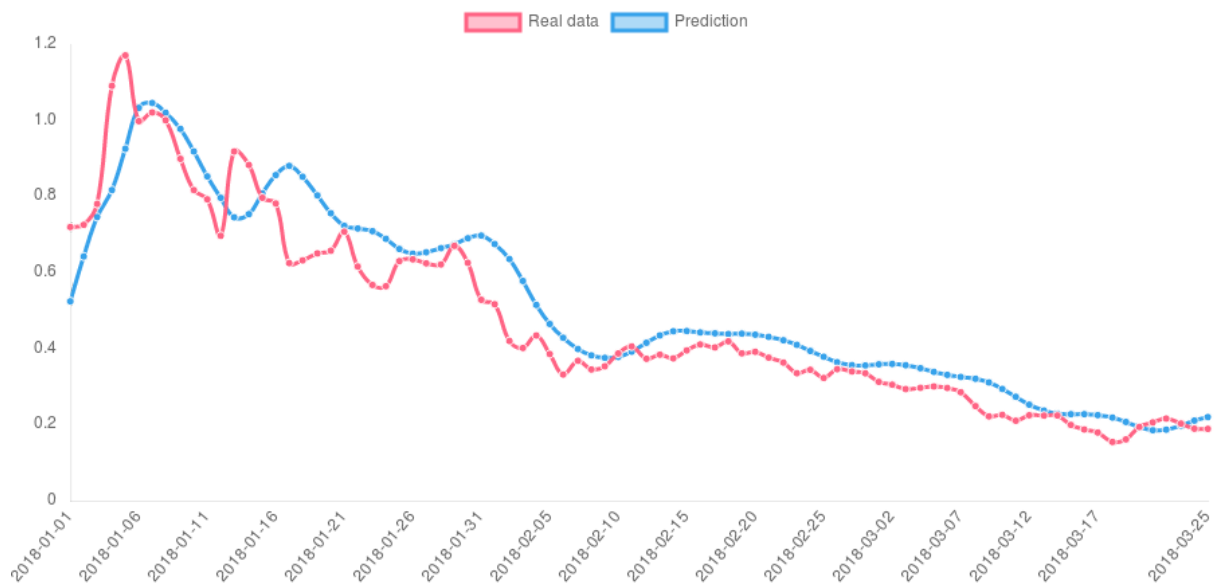
6.g Prediccions a valor únic bitCoin Cash (R^2 de 90%) [creació pròpia]



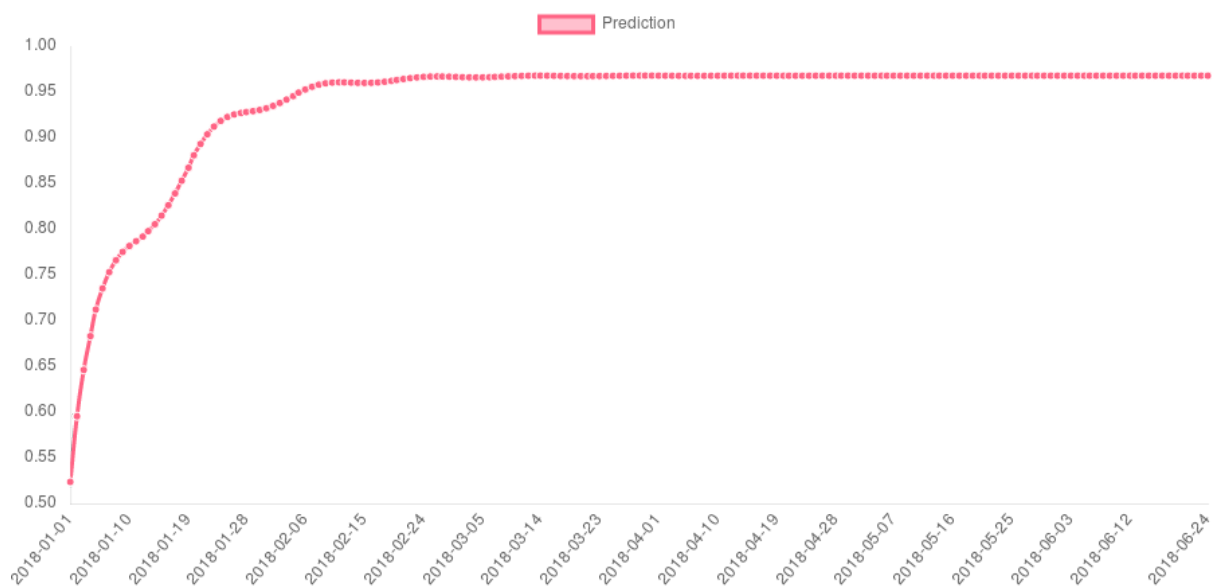
6.h Prediccions futures bitCoin Cash (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica 6.g podem observar que el marge d'error entre els valors reals i els valors predits no supera els 300 \$ per a valors que es mouen entre els 1000 \$ i els 3000 \$. A la gràfica 6.h es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 5800 \$. Per últim es pot veure que l' R^2 d'aquesta predicció és del 90%.

Cardano



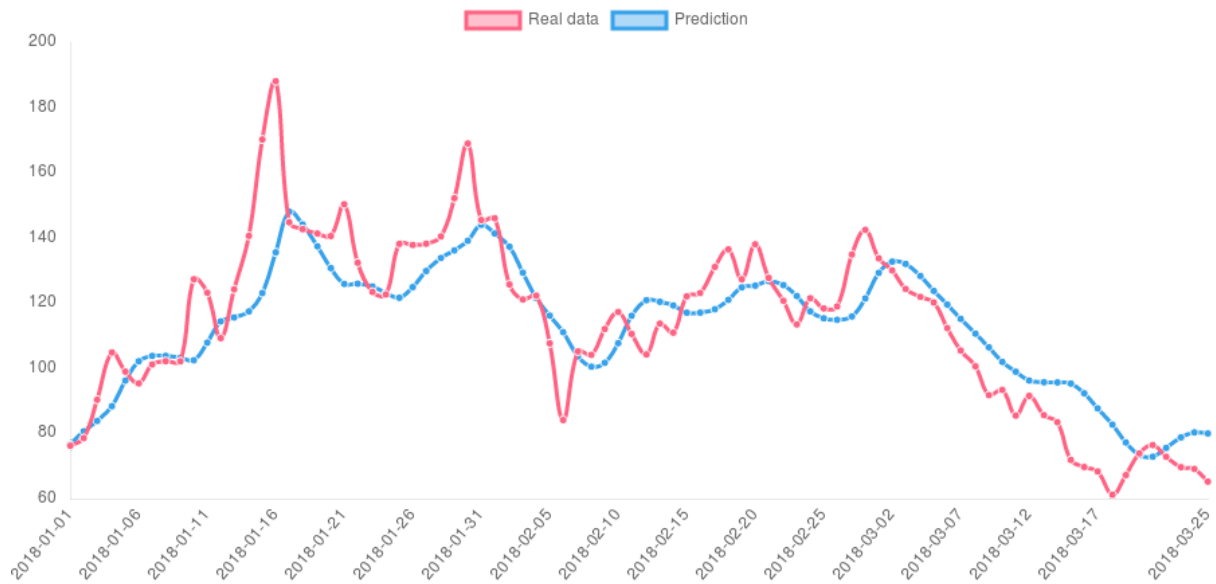
6.i Prediccions a valor únic Cardano (R^2 de 89%) [creació pròpia]



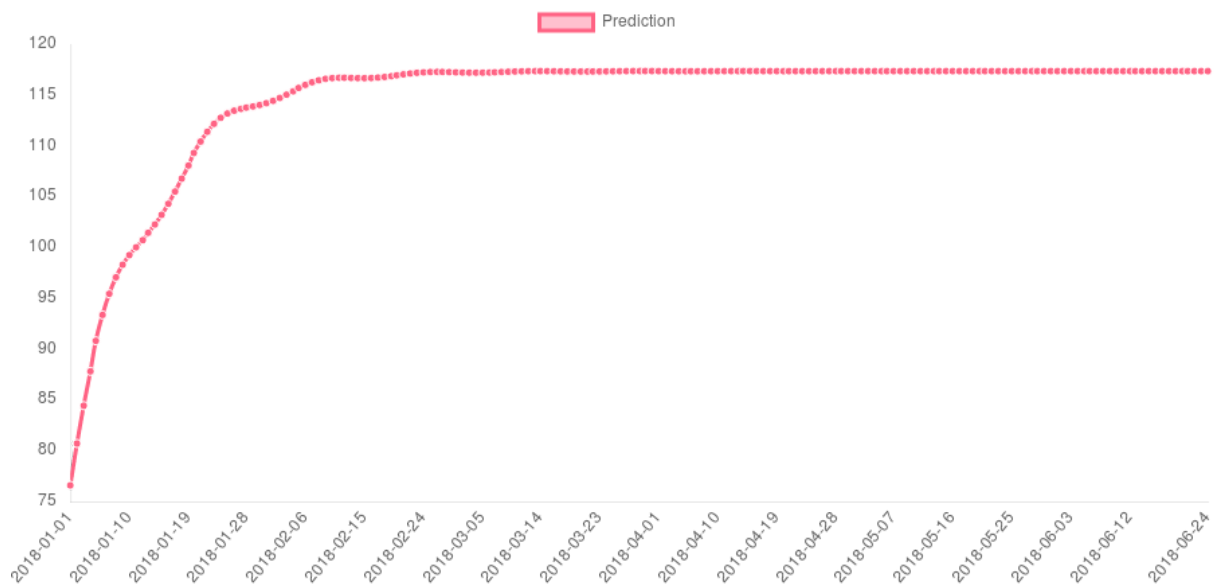
6.j Prediccions futures Cardano (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica 6.i podem observar que el marge d'error entre els valors reals i els valors predits no supera els 0.3 \$ per a valors que es mouen entre els 0.2 \$ i els 1.2 \$. A la gràfica 6.j es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 0.97 \$. Per últim es pot veure que l' R^2 d'aquesta predicció és del 89%.

NEO



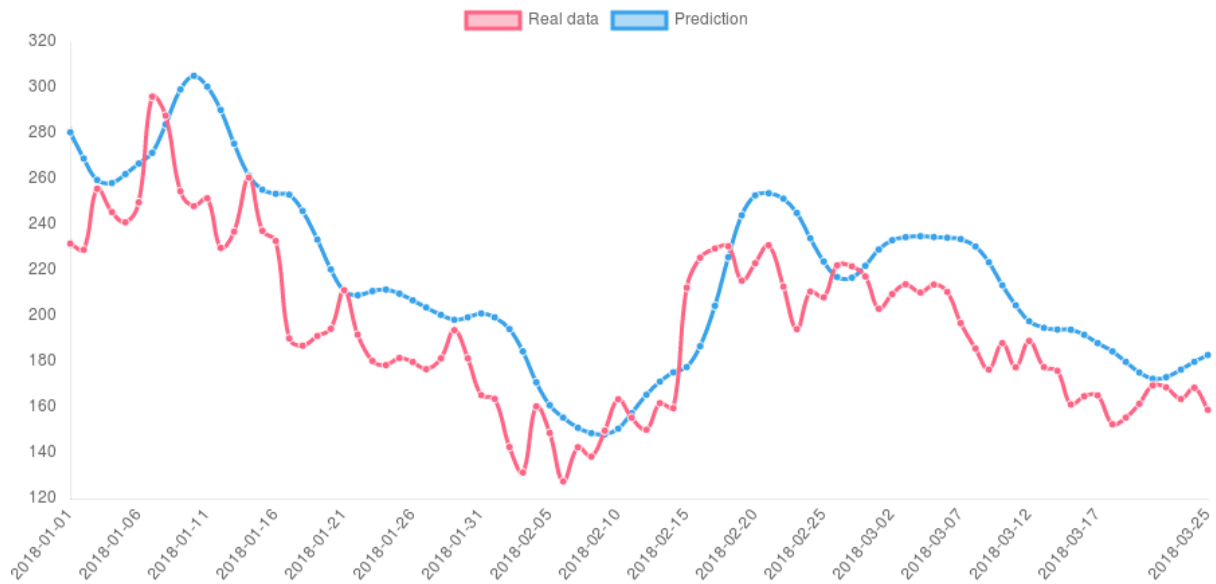
6.k Prediccions a valor únic NEO (R^2 de 76%) [creació pròpia]



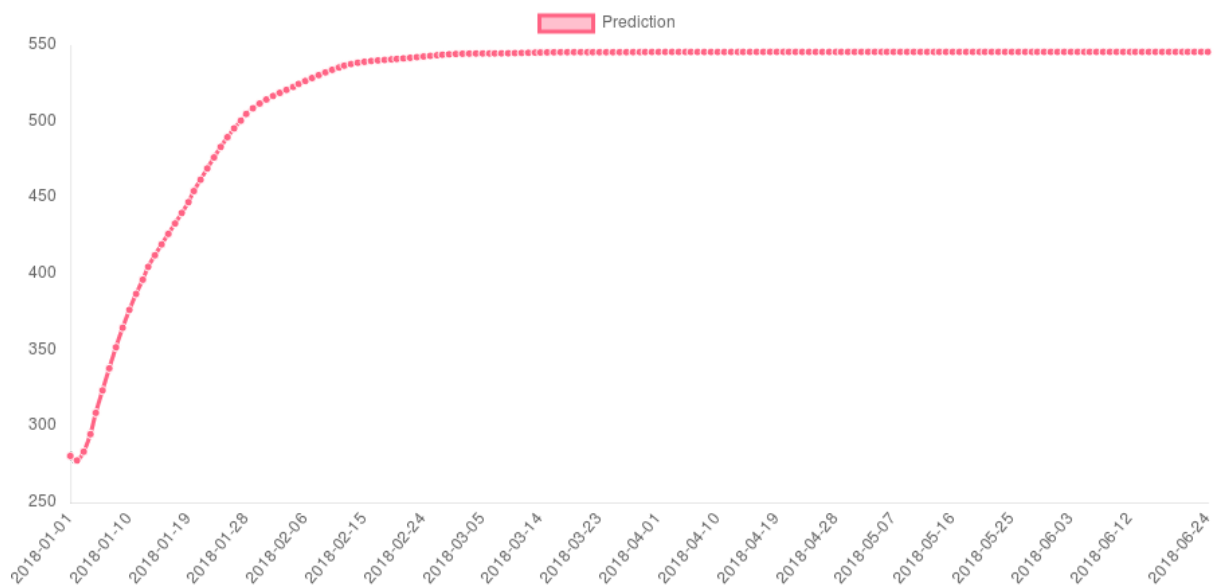
6.l Prediccions futures NEO (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica 6.k podem observar que el marge d'error entre els valors reals i els valors predits no supera els 50 \$ per a valors que es mouen entre els 60 \$ i els 190 \$. A la gràfica 6.l es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 117 \$. Per últim es pot veure que l' R^2 d'aquesta predicció és del 76%.

Litecoin



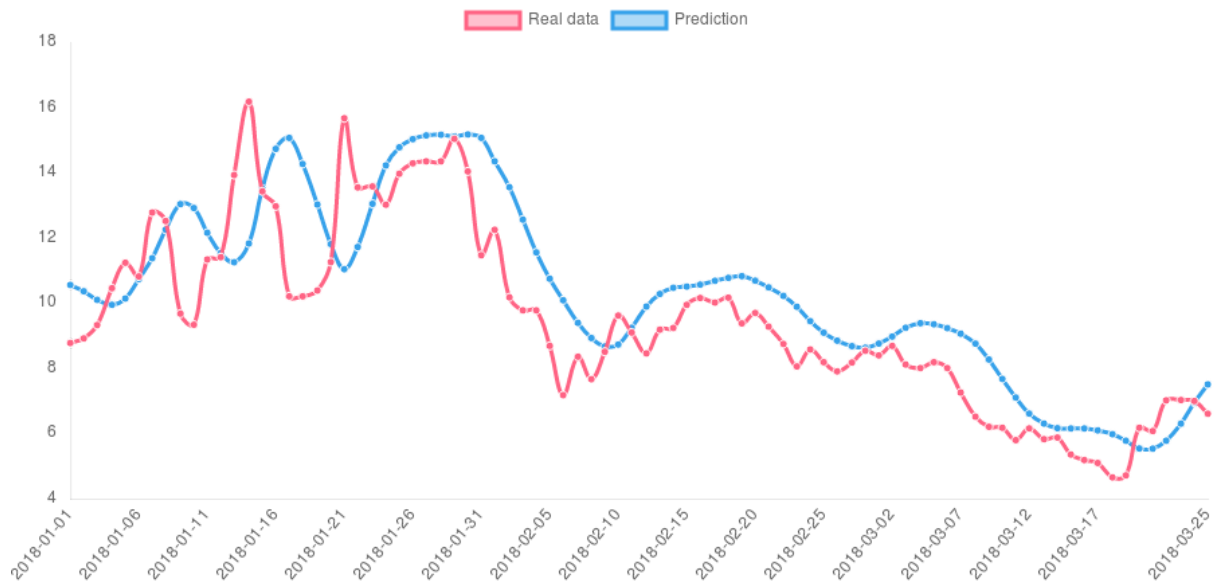
6.m Prediccions a valor únic Litecoin (R^2 de 73%) [creació pròpia]



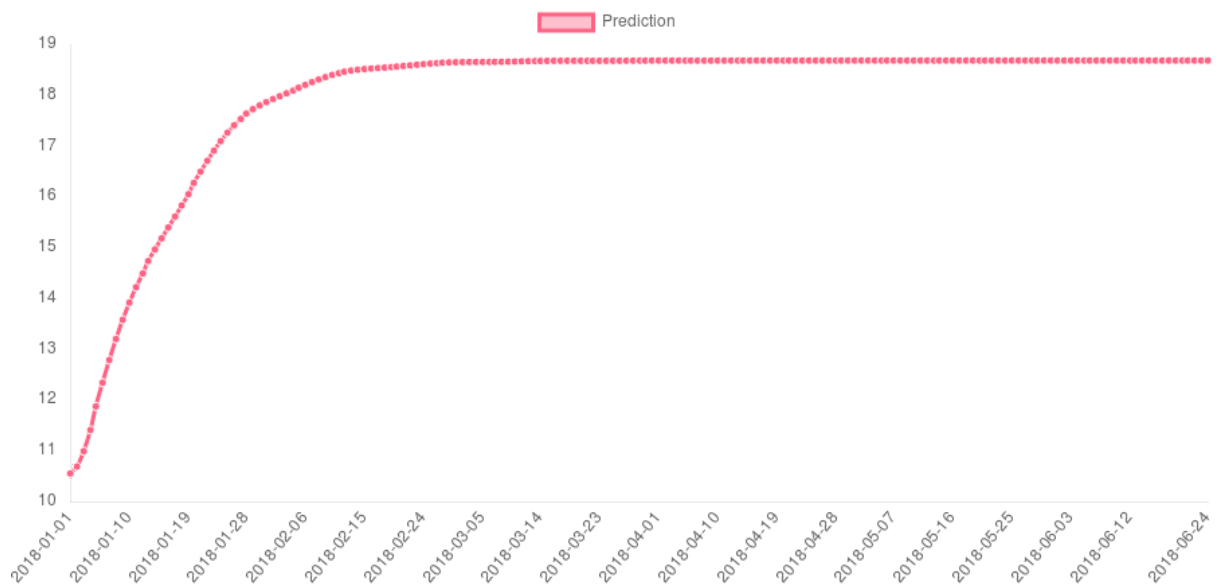
6.n Prediccions futures Litecoin (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica **6.m** podem observar que el marge d'error entre els valors reals i els valors predits no supera els 60 \$ per a valors que es mouen entre els 120 \$ i els 300 \$. A la gràfica **6.n** es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 540 \$. Per últim es pot veure que l' R^2 d'aquesta predicció és del 73%.

EOS



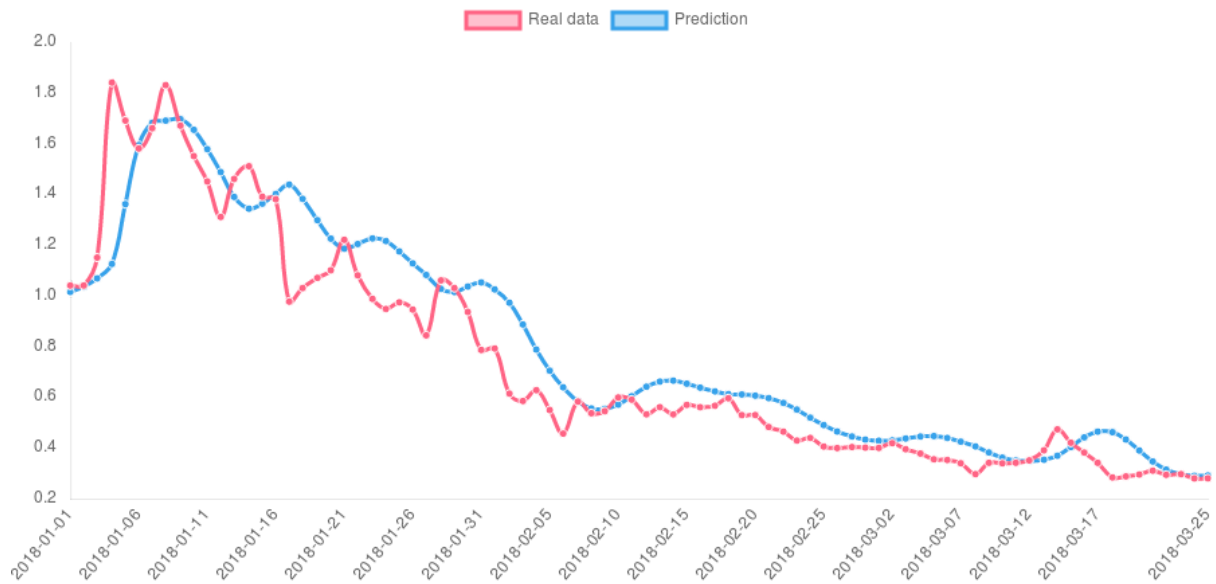
6.o Prediccions a valor únic EOS (R² de 72%) [creació pròpia]



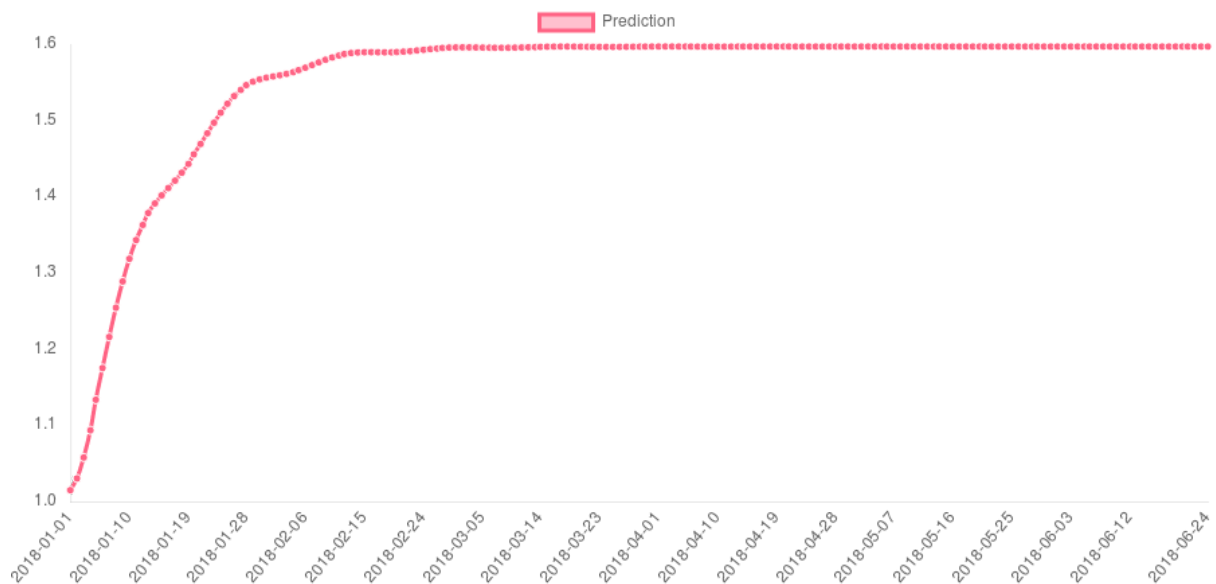
6.p Prediccions futures EOS (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica 6.o podem observar que el marge d'error entre els valors reals i els valors predits no supera els 5 \$ per a valors que es mouen entre els 5 \$ i els 17 \$. A la gràfica 6.p es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 18 \$. Per últim es pot veure que l'R² d'aquesta predicció és del 72%.

NEM



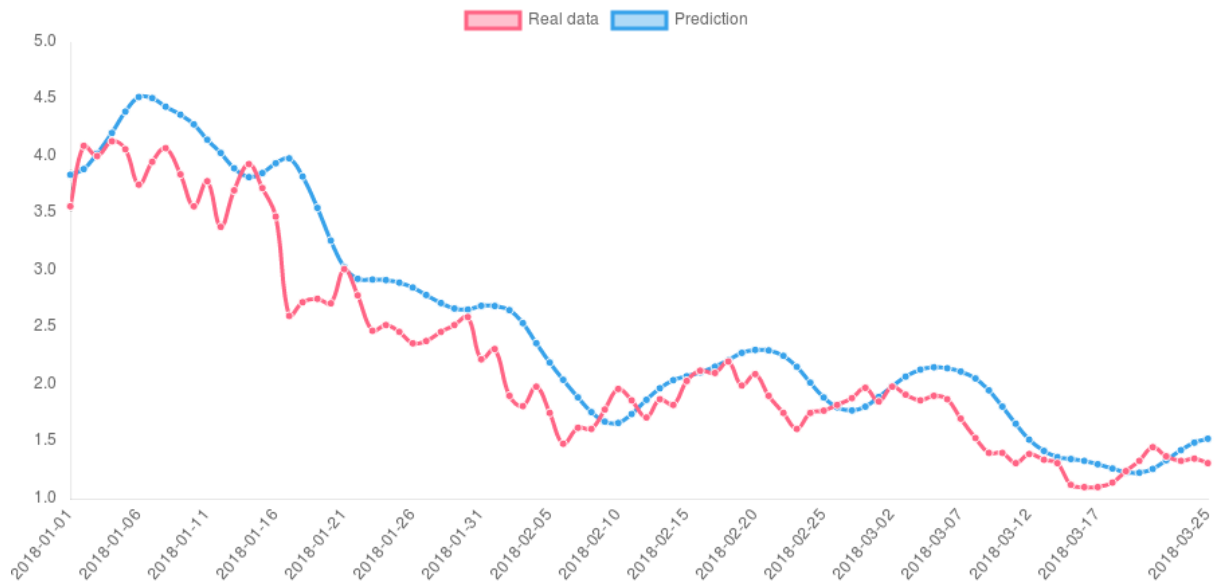
6.q Prediccions a valor únic NEM (R^2 de 89%) [creació pròpia]



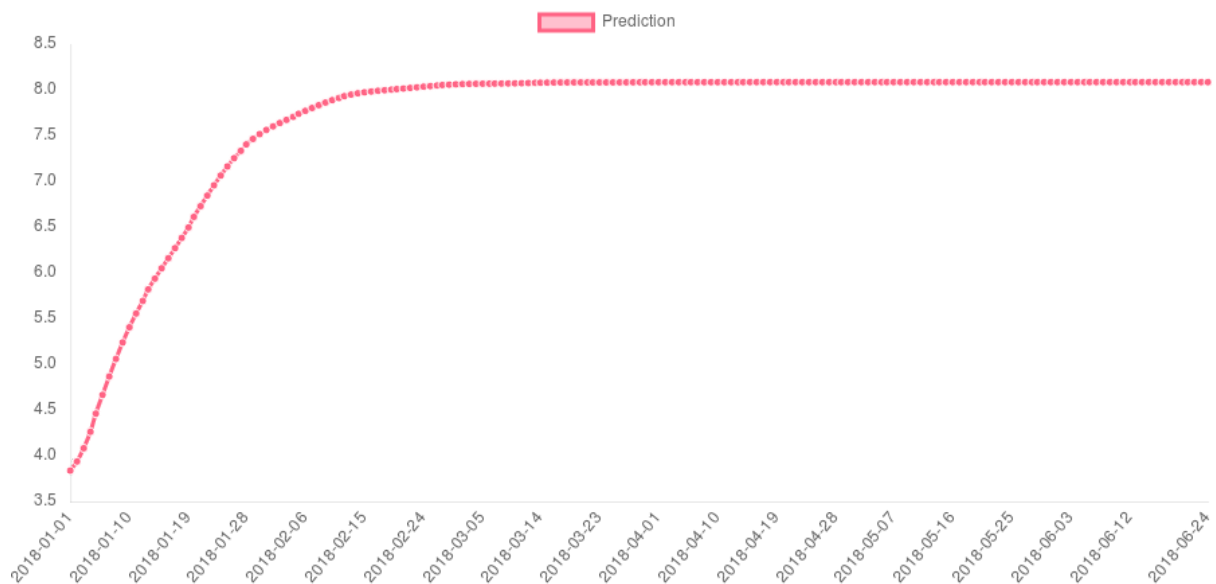
6.r Prediccions futures NEM (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica **6.q** podem observar que el marge d'error entre els valors reals i els valors predits no supera els 0.2 \$ per a valors que es mouen entre els 0.3 \$ i els 1.8 \$. A la gràfica **6.r** es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 1.6 \$. Per últim es pot veure que l' R^2 d'aquesta predicció és del 89%.

IOTA



6.s Prediccions a valor únic IOTA (R² de 91%) [creació pròpia]



6.t Prediccions futures IOTA (fins el 25 de juny de 2018) [creació pròpia]

En la gràfica 6.s podem observar que el marge d'error entre els valors reals i els valors predits no supera els 1 \$ per a valors que es mouen entre els 1 \$ i els 4.1 \$. A la gràfica 6.t es pot observar que l'error acumulat de les prediccions fa que s'estableixi un límit de valors en 8 \$. Per últim es pot veure que l'R² d'aquesta predicció és del 91%.

7. Autoavaluació del domini actual de la competència de sostenibilitat.

Per començar, faré una anàlisi sobre el coneixement actual que tinc sobre la competència de sostenibilitat basant-nos en l'enquesta que se'ns ha demanat [17].

A gran trets, l'enquesta avalua tres aspectes sobre les sostenibilitats ambiental, econòmica i social: nivell de coneixement de les problemàtiques de l'àrea, capacitat d'identificar-les en els projectes i capacitat de proposar solucions en els projectes que les tinguin en compte.

Sobre el nivell de coneixement de les problemàtiques, crec que, tot i que no és el més avançat, hem adquirit els coneixements suficients al llarg de les diferents matèries del grau. Sobretot en casos com el meu on he participat en assignatures optatives com ASMI (Aspectes Socials i Mediambientals de la Informàtica) que tracta en gran mesura aquests aspectes.

On crec que hi ha la meva major manca de coneixements és en l'identificació de les problemàtiques relacionades amb la sostenibilitat en els projectes. Això pot ser degut a que durant els cursos de grau on s'ha tractat se'ns ha fet analitzar de manera teòrica les implicacions d'una implementació però ha sigut poc habitual que se'ns demani que afegim identificadors de sostenibilitat per tal d'analitzar-los. És interessant aquesta enquesta ja que ens fa pensar i avaluar aquests aspectes que normalment no tenim en compte.

Per últim, crec que el meu punt més fort és la capacitat de proposar solucions sostenibles per als reptes. Tot i que a primera vista no té gaire sentit pel comentat en el punt anterior, si ho analitzem detingudament sí que el té. Durant el grau aprenem a ser capaços de trobar solucions als problemes que apareixen a diari i per tant, si abans de realitzar un projecte ens parem a analitzar la seva sostenibilitat hauríem de ser capaços de trobar aquelles solucions que millor s'hi adaptin.

A mode de conclusió, puc dir que el meu punt feble és dissenyar identificadors de sostenibilitat en els projectes amb l'objectiu que indiquin el grau de sostenibilitat que té i el meu punt fort és trobar solucions sostenibles a les problemàtiques dels projectes.

8. Pressupost

Una vegada planificat el projecte, es va fer un estudi dels costos que comportaria la realització del mateix per poder fer-ne un pressupost i planificar les despeses. També es van analitzar les possibles desviacions que podien sorgir en el pressupost. Al llarg d'aquest capítol podem veure la comparativa entre els pressuposts planificats i les despeses reals.

8.1. Estimació de costos

Per poder realitzar una estimació adequada es van definir els elements següents:

Costos directes

A continuació podem observar una taula que mostra les hores dedicades per cada rol i el cost per hora segons la planificació inicial. No s'afegeixen costos de software ja que tots són gratuïts.

Rol	Hores	Cost/hora [18]	Total
Cap de projecte	144	28,75 €	4140 €
Analista	188	16,75 €	3149 €
Programador	140	15,00 €	2100 €
Provador	40	15,00 €	600 €
Total			9989 €

8.1.a. Estimació de costos directes segons cada rol. [creació pròpia]

A continuació podem observar una taula que mostra les hores reals dedicades per cada rol i el cost per hora. No s'afegeixen costos de software ja que tots són gratuïts.

Rol	Hores	Cost/hora [18]	Total
Cap de projecte	145	28,75 €	4170 €
Analista	139	16,75 €	2330 €
Programador	194	15,00 €	2910 €
Provador	50	15,00 €	750 €
Total			10160 €

8.1.b. Costos directes reals segons cada rol. [creació pròpia]

Costos indirectes

A continuació veurem la taula que mostra les despeses generades indirectament pel projecte segons la planificació inicial.

Nom	Temps	Cost	Total
Electricitat	512 hores	0,17 / hora €	87,04 €
Transport (T-Jove) [19]	3 mesos	105 €	105 €
Servidor [20]	3 mesos	16,28 € / mes	48,84 €
Internet + telèfon [21]	3 mesos	33,06 € / mes	99,18 €
Aigua	3 mesos	18,28 € / mes	54,84 €
Lloguer + assegurança	3 mesos	800 € / mes	2400 €
Total			3035,78 €

8.1.c. Estimació de costos indirectes segons cada recurs. [creació pròpia]

A continuació veurem la taula que mostra les despeses reals generades indirectament pel projecte.

Nom	Temps	Cost	Total
Electricitat	528 hores	0,17 / hora €	89,76 €
Transport (T-Jove) [19]	3 mesos	105 €	105 €
Servidor [20]	0	16,28 € / mes	0 €
Internet + telèfon [21]	3 mesos	33,06 € / mes	99,18 €
Aigua	3 mesos	18,28 € / mes	54,84 €
Lloguer + assegurança	3 mesos	800 € / mes	2400 €
Total			2748,78 €

8.1.d. Costos indirectes reals segons cada recurs. [creació pròpia]

Amortitzacions

A continuació avaluem les diferents reduccions en els valors dels productes segons les amortitzacions. L'amortització relativa fa referència als mesos que durarà el projecte que serà un trimestre ($33\% / 4 = 8,25\%$):

Nom	Preu (€)	Amortització anual	Amortització relativa	Amortització (€)
Portàtil Hp pavilion [22]	999,90	33%	8,25%	83,32
Monitor Dell [23]	150,04	33%	8,25%	12,50
Total	1140,94			95,82

8.1.e. Estimació d'amortitzacions. [creació pròpia]

En el cas de les amortitzacions el cost real i el cost planificat has estat els mateixos.

8.2. Control de gestió

Els valors calculats fins ara no permetien cap desviació ni cap error, ja que es suposa que tot sortirà exactament com s'ha planejat. Si el projecte s'acabés abans del previst, la resta d'hores planificades es dedicarien a millorar la feina realitzada o a incloure noves funcionalitats. Altrament, si el projecte necessités més hores trobaríem que el pressupost calculat no és suficient.

Desviació del temps

Com hem comentat anteriorment, si en l'última iteració del projecte no s'assoleixen els objectius especificats, es consideraria dedicar més hores al desenvolupament d'aquest. Per tenir un cert marge en el pressupost es van calcular que aquestes hores addicionals podien arribar a ser fins un 10% de les hores planificades. Això resultaria en un augment d'unes 51 hores com a màxim que farien afegir al pressupost **768 euros**.

Avaria en l'ordinador o pantalla

Un dels possibles imprevistos que podien sorgir i afectar al pressupost era una avaria en el portàtil on es realitzava el projecte. En aquest cas, s'hauria d'haver comprat un de nou i recalculat l'amortització de nou. Suposant que aquest imprevist hagués aparegut a meitat de la planificació, hagués implicat una amortització de un mes i mig que es traduïa en **41,66 euros** addicionals. En cas que l'avaria s'hagués produït en la pantalla el cost addicional hagués estat **6,25 euros**.

Contingència

Per últim, per assegurar cobrir altres despeses que no s'haguessin pogut planificar, es va aplicar un percentatge de contingències al cost total del projecte, així s'assegurava un compliment molt més fiable del pressupost que en cap cas es superaria. Es va decidir que el marge de contingències fos del **5%**.

8.3. Pressupost final

Un cop s'havien tingut en compte tots els costos i imprevistos, el pressupost final planificat inicialment que vam obtenir va ser:

Nom	Cost (€)
Costos directes	9989
Costos indirectes	3035,7
Amortitzacions	95,82
Imprevistos	774,25
Contingències	695
Total	14589,77

8.3.a. Estimació total dels costos. [creació pròpia]

Les despeses totals reals del projecte han estat:

Nom	Cost (€)
Costos directes	10160
Costos indirectes	2748,78
Amortitzacions	95,82
Imprevistos	774,25
Contingències	695
Total	14473.85

8.3.b. Total dels costos reals. [creació pròpia]

9. Informe de sostenibilitat

En aquest capítol es fa una reflexió sobre com el nostre projecte ha afrontat la problemàtica de la sostenibilitat. Per fer-ho, s'analitza la matriu de sostenibilitat.

9.1. Estudi d'impacte ambiental

En aquesta secció es realitza un estudi sobre les conseqüències ambientals que pot suposar aquest projecte.

Durant la fase de planificació del projecte es va realitzar una estimació sobre l'impacte ambiental del treball. Per a aquesta estimació es va argumentar que es reutilitzarien servidors propis de l'empresa per tal de no tenir que afegir-ne de nous. A la taula següent podem observar l'estimació de l'impacte ambiental:

Nom	Hores	Cost/hora (kW/hora)	Total (kW)
1 Persona en rutina habitual	512	0,1	51,2
Pantalla	512	0,03	15,36
Portàtil	512	0,125	64
Servidor	512	0,125	64
Total			194,56

9.1.a. Estimació de l'impacte ambiental. [creació pròpia]

Durant la realització del projecte es va decidir no utilitzar un servidor i mantenir tot el desenvolupament en local per tal d'evitar fer ús i així reduir l'impacte ambiental. El resultat d'aquesta decisió es veu reflectit en la taula real d'impacte ambiental:

Nom	Hores	Cost/hora (kW/hora)	Total (kW)
1 Persona en rutina habitual	528	0,1	52,8
Pantalla	528	0,03	15,84
Portàtil	528	0,125	66
Servidor	0	0,125	0
Total			134,64

9.1.b. Impacte ambiental real. [creació pròpia]

En cas de realitzar el projecte de nou creiem que l'únic recurs del que es pot prescindir es la pantalla perquè només amb la del portàtil es pot treballar encara que sigui més incòmode fer-ho amb una de sola.

Avui dia la majoria de reptes basats en mercats financers es resolen mitjançant regressions lineals o múltiples. Per implementar aquestes tècniques s'utilitza software que, tot i que és un software diferent al utilitzat en aquest projecte, ambientalment tenen el mateix impacte. Per tant, la solució utilitzada per a la realització del projecte no aporta cap millora envers les altres disponibles.

Durant la vida útil del projecte s'utilitzarà un servidor petit per tal de donar accés al públic. Aquest servidor serà una instància EC2 d'Amazon, el motiu d'aquesta elecció es fonamenta en el compromís d'Amazon amb la sostenibilitat [24]. No obstant, si el nombre d'usuaris que accedeixen escala ràpid s'haurà de pensar en solucions més complexes en Amazon. Això faria augmentar el nombre de recursos a utilitzar i la petjada ecològica es veuria perjudicada. En canvi, si l'ús fos inexistent, Amazon permet eliminar les instàncies creades, eliminant així qualsevol ús de recursos.

9.2. Estudi d'impacte econòmic

Al capítol 8 s'expliquen els recursos que es van estimar inicialment que utilitzaríem i els recursos que al final s'han utilitzat en el desenvolupament. Les xifres de l'àmbit dels recursos humans estan pensades per a un projecte real, però tractant-se d'un TFG el projecte ha estat molt més econòmic que no pas si l'hagués fet una empresa pel seu compte ja que qui ha realitzat tot el disseny, planificació i implementació ha estat un becari que costa 8 euros l'hora. Aquest fet ens estalvia al voltant de 5936 € ja que els sous dels diferents rols són més elevats. No obstant, tant els pressuposts estimats com els reals s'han fet basant-se en rols, per tant, aquesta quantitat no es comptarà com a estalvi.

Una de les decisions que es van prendre per minimitzar l'impacte econòmic va ser no utilitzar cap servidor i realitzar tota la implementació en entorns locals. Aquest fet ens estalvia al voltant de 50 €.

El pressupost estimat s'ha ajustat bastant als valors reals, la diferència ha estat menor que la quantitat que s'havia destinat per a imprevistos i com a contingència. És a dir, el nostre pla d'acció contra imprevistos ha funcionat i no ens hem sortit de l'estimació.

Avui dia la majoria de reptes basats en mercats financers es resolen mitjançant regressions lineals o múltiples. Per implementar aquestes tècniques s'utilitzen softwares gratuïts que són més fàcils d'entendre i d'implementar que no pas el software que hem fet servir en aquest projecte. En l'àmbit econòmic s'han hagut de dedicar més hores en programar que no pas si haguéssim fet cap altra solució. Aquestes hores es tradueixen en valors econòmics i per tant, realitzar aquest projecte ha estat més car en comparació a realitzar una solució basada en regressions. La avantatge que tenim econòmicament respecte les altres solucions és el coneixement adquirit, que permetrà, a partir d'ara, oferir un nou servei als clients de l'empresa, un servei de xarxes neuronals com a eina per resoldre reptes.

Durant la vida útil del projecte estimem que el cost que tindrà serà el preu que ens demani Amazon per l'ús de la seva instància EC2. Aquest cost inicialment estarà al voltant dels 20 €/mes. No obstant, si l'ús comença a escalar ràpidament s'haurà de pensar en solucions més complexes d'Amazon, que faria augmentar el nombre de recursos que s'utilitzarien i també els costos. En canvi, si ningú fa ús del resultat del projecte es pot eliminar l'instància d'Amazon eliminant també el cost d'aquesta.

9.3. Estudi d'impacte social

Durant la realització del projecte s'han adquirit un gran nombre de coneixements sobre què són les xarxes neuronals, com funcionen, com es comuniquen les neurones, quins tipus de xarxes existeixen, ... Aquests coneixements són la principal aportació que es rep del desenvolupament del projecte.

La realització del projecte ha fet que es plantegin certes sortides professionals més enfocades al Machine Learning i el Deep Learning. Aquestes sortides són cada cop més demandades en el món de la intel·ligència artificial.

La solució al problema base de les criptomonedes no aporta cap millora de qualitat de vida a la societat. Però haver adquirit tot el coneixement que hem adquirit obre portes per oferir serveis que millorin en un futur la qualitat de vida de la societat. Per altra banda, aquest projecte satisfà la necessitat de SlashMobility d'obrir noves vies de servei als seus clients, que es beneficiaran d'aquest fet.

El projecte podria generar alguna dependència als addictes d'inversions ja que si confien el suficient en els resultats poden arribar a fer grans inversions en criptomonedes que després perdin si les prediccions no son correctes.

La solució implementada en aquest projecte compleix alguns dels reptes proposats. Per una banda compleix l'objectiu d'aprendre el suficient de Deep Learning com per a oferir nous serveis als clients. Per altra banda, les prediccions són parcialment satisfactòries ja que en algunes de les criptomonedes s'han obtingut R^2 elevats. El problema principal d'aquestes prediccions és que es realitzen sobre valors únics, és a dir, sempre i quan tots els valors anteriors siguin informació real s'obté l' R^2 esmentat. Però en el cas que les prediccions es realitzin sobre valors resultats d'altres prediccions s'obtenen valors massa erronis ja que s'acumulen els marges d'error de les prediccions anteriors.

9.4. Conclusions

En els estudis de les seccions anteriors d'aquest capítol s'ha realitzat una anàlisi basada en la matriu de sostenibilitat. Per acabar amb aquest capítol, a continuació es puntuen els resultats i s'obtenen els totals per a cada àmbit:

	PPP	Vida útil	Riscos	Total
--	-----	-----------	--------	-------

Ambiental	10	15	-5	20
Econòmic	8	15	-5	18
Social	5	10	-10	5
Total	23	40	-20	43

9.4 Taula de matriu de sostenibilitat. [creació pròpia]

10. Lleis i regulacions

En aquest capítol tractarem d'identificar si el nostre projecte es veu regulat per alguna llei que ens faci modificar o adaptar certs aspectes del treball.

Una de les lleis més comunes en els projectes software és la protecció de dades, i ara és més important si cap amb la entrada en vigor de la General Data Protection Regulation (GDPR) [25]. Aquestes normes busquen que les autoritzacions dels usuaris per al tractament de dades siguin explícites i s'obligui als usuaris a conèixer de manera més entenedora com es tractaran les seves dades. En tot cas, en el nostre projecte no hi ha cap registre d'usuaris, i ni tan sols s'emmagatzemen dades, per tant, no s'ha de prendre cap mesura per satisfer aquesta regulació.

Pel que fa a les llicències, tot el software que hem utilitzat és lliure i no és necessari pagar cap llicència ni seguir cap regulació. Les dades de les criptomonedes són públiques i per tant són de lliure utilització.

11. Conclusions

Com a conclusions del projecte podem dir que s'ha estudiat el funcionament de les xarxes neuronals, adquirint coneixement sobre els algoritmes que hi han darrere d'aquestes i que fan possible les comunicacions entre neurones, entre d'altres coses. També s'ha après que existeixen diferents tipus de xarxes neuronals i hem entrat en detall en les xarxes neuronals recurrents per resoldre el problema plantejat a les fases inicials del projecte.

Pel que fa als frameworks que proporcionen mètodes per implementar xarxes neuronals s'han estudiat diverses possibilitats, entrant en detall en el framework Keras que gràcies a la seva capacitat de treballar per sobre d'altres frameworks com Tensorflow proporciona una gran quantitat d'informació sobre com desenvolupar xarxes neuronals.

Com a conclusions sobre els resultats de la configuració i entrenament de les xarxes neuronals podem dir que per a tot problema es necessita provar diferents combinacions de paràmetres per tal d'obtenir la combinació més òptima. Per fer-ho el propi framework de Keras ens proporciona un avaluador configurable. Tot i així, hem après què significa cada paràmetre i a raonar el perquè dels resultats obtinguts en aquest aspecte.

Com es pot observar en totes les criptomonedes en les que s'ha provat la xarxa neuronal, els resultats en valor únic són molt positius ja que totes elles han superat l'objectiu inicial de un R^2 superior al 70%. Però hi ha un problema quan intentem realitzar prediccions futures perquè l'error de cada dia predit s'acumula per a les prediccions següents, fent que arribat cert punt d'error acumulat les prediccions s'estabilitzin generant tendència a un límit.

Una observació important que es pot realitzar es l'adaptabilitat de la xarxa neuronal. Per una banda, quan els canvis dels valors són lineals la xarxa prediu molt acuradament generant errors molts petits, i això és degut a que el model ha après a predir el que hauria de passar en situacions normals. Per altra banda, quan apareix una singularitat que implica una baixada o una pujada ràpida de valors a la xarxa li costa més ja que no és habitual que això passi. Aquestes singularitats acostumen a aparèixer per factors externs i és per això que la xarxa no els detecta. A més a més, d'acord amb el concepte matemàtic de Brownian motion en enginyeria financera, les versions futures dels valors dels preus de les criptomonedes són independents del passat i, per tant, aquestes singularitats que la xarxa no detecta adequadament són independents del que hagi ocorregut en el passat.

Els futurs passos que es poden fer per continuar millorant la xarxa neuronal són:

- Continuar afegint les dades dia a dia com a entrenament de la xarxa.
- Afegir dades de factors externs com poden ser notícies importants en l'àmbit financer o criptofinancer. Per a aquesta part s'haurà d'afegir una nova xarxa neuronal que sigui capaç d'identificar el grau d'afectació de les notícies en funció del contingut d'aquestes i afegir aquesta nova variable a la xarxa actual.

A mode personal, la realització d'aquest projecte ha estat molt satisfactòria per diversos motius. Per una banda, he après moltíssim sobre Machine Learning i Deep Learning arribant a programar una xarxa neuronal funcional que és capaç de predir bastant acuradament. Per altra banda he descobert tota una àrea d'investigació important que m'ha atret des del primer moment. En quant al resultat del projecte m'hagués agradat haver-lo dut un pas més enllà i haver inclòs dades de factors externs que afectessin els resultats de les prediccions. No obstant, aquest projecte continuarà creixent en un futur ja sigui amb l'empresa o per motivació personal.

A nivell d'empresa, Slashmobility ha obtingut un treballador amb coneixement sobre xarxes neuronals que permet a l'empresa oferir un nou servei als seus clients, un servei de resolució de reptes mitjançant xarxes neuronals. A banda d'això, Slashmobility està cada cop entrant més en àrees d'innovació i el software de prediccions ajuda en les xerrades tècniques que es realitzen en diferents llocs per divulgar coneixements sobre Blockchain i xarxes neuronals.

11. Referències

- [1] - SlashMobility (2018). *SlashMobility* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <http://slashmobility.com/> >
- [2] - Michael Nielsen (2017). *Neural networks and deep Learning* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <http://neuralnetworksanddeeplearning.com/about.html> >
- [3] - Kirill Eremenko (2018). *Kirill Eremenko* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <https://au.linkedin.com/in/keremenko> >
- [4] - Hadelin de Ponteves (2018). *Hadelin de Ponteves* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <https://www.linkedin.com/in/hadelin-de-ponteves-1425ba5b> >
- [5] - Kirill Eremenko & Hadelin de Ponteves (2017). *Machine Learning A-Z™* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <https://www.udemy.com/machinelearning/learn/v4/overview> >
- [6] - MathWords (2017). *Independent Variable* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < http://www.mathwords.com/i/independent_variable.htm >
- [7] - MathWords (2017). *Dependent Variable* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < http://www.mathwords.com/d/dependent_variable.htm >
- [8] - Phylliida (2016). *A list of cost functions used in neural networks, alongside applications* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <https://stats.stackexchange.com/questions/154879/a-list-of-cost-functions-used-in-neural-networks-alongside-applications> >
- [9] - iamtrask (2015). *A Neural Network in 13 lines of Python* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <http://iamtrask.github.io/2015/07/27/python-network-part2/> >
- [10] - Michael Nielsen (2017). *Neural networks and deep Learning* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <http://neuralnetworksanddeeplearning.com/chap2.html> >
- [11] - Yann LeCun, Leon Bottou, Genevieve B. Orr & Klaus-Robert Müller (1998). *Efficient BackProp* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <http://yann.lecun.com/exdb/publis/pdf/lecun-98b.pdf> >
- [12] - TensorFlow (2018). *TensorFlow* [en línia]. [Consultada: 4 de març 2018]. Disponible a Internet: < <https://www.tensorflow.org/> >

[13] - Theano (2018). *Theano* [en línea]. [Consultada: 4 de març 2018]. Disponible a Internet: < <http://deeplearning.net/software/theano/> >

[14] - Keras (2018). *Keras* [en línea]. [Consultada: 4 de març 2018]. Disponible a Internet: < <https://keras.io/> >

[15] - Torch (2018). *Torch* [en línea]. [Consultada: 4 de març 2018]. Disponible a Internet: < <http://torch.ch/> >

[16] - Caffe (2018). *Caffe* [en línea]. [Consultada: 4 de març 2018]. Disponible a Internet: < <http://caffe.berkeleyvision.org/> >

[17] - Google Form (2018). *Google Form* [en línea]. [Consultada: 20 de març 2018]. Disponible a Internet: < <goo.gl/kWLMLE> >

[18] - Average Salary Survey (2017). *Average Salary Survey* [en línea]. [Consultada: 8 d'abril 2018]. Disponible a Internet: < <http://www.averagesalariesurvey.com/> >

[19] - TMB (2017). *T-Jove* [en línea]. [Consultada: 22 de març 2018]. Disponible a Internet: < <https://www.tmb.cat/es/barcelona/tarifas-metro-bus/abonos/t-jove> >

[20] - Digital Ocean (2017). *Droplets* [en línea]. [Consultada: 8 d'abril 2018]. Disponible a Internet: < <https://www.digitalocean.com/pricing/> >

[21] - Movistar (2017). *Fibra 50 MB* [en línea]. [Consultada: 22 de març 2018]. Disponible a Internet: < <http://www.movistar.es/empresas/>>

[22] - FNAC (2017). *Portàtil Hp Pavilion* [en línea]. [Consultada: 22 de març 2018]. Disponible a Internet: < <https://www.fnac.es/Portatil-HP-Pavilion-Notebook-14-bf101ns-Plata-Ordenador-portatil/a1400596> >

[23] - DELL (2017). *Monitor Dell* [en línea]. [Consultada: 22 de març 2018]. Disponible a Internet: < <http://www.dell.com/es-es/shop/monitor-dell-24-se2416h/apd/210-afzc/monitores-y-accesorios> >

[24] - Amazon (2016). *Amazon Sostenibilidad* [en línea]. [Consultada: 20 de juny 2018]. Disponible a Internet: < <https://www.amazon.es/p/feature/xm33ycveeesj454> >

[25] - GDPR (2016). *General Data Protection Regulation* [en línea]. [Consultada: 23 de maig 2018]. Disponible a Internet: < <https://gdpr-info.eu/> >

