

PLAY & LEARN:

Eina de suport a l'aprenentatge combinant estratègies de gamificació i Machine learning

Oriol Catalan Yarza

Resum—Creació d'un mòdul de codi amb tècniques d'aprenentatge computacional per al processament del llenguatge i classificació, per a la seva incorporació en el desenvolupament de videojocs enfocats a l'educació, emprant recomanació de preguntes al jugador basades en preguntes fallades anteriorment. S'utilitzaran les mètriques de *Accuracy*, *Precision*, *Recall*, i *Accuracy per class* mitjançant la creació d'una *matriu de confusió*. També es comentaran els resultats obtinguts així com les diferents estratègies que es poden aplicar per a millorar aquests resultats obtinguts.

Paraules clau—Aprenentatge Computacional, processament del llenguatge, classificació, histograma, Accuracy, Precision, Recall, Accuracy per class, Matriu de confusió, Unity.

Abstract—Creation of a code module with Machine Learning techniques for Language procesing and classification, to be incorporated in videogame develepment focused on education, by means of recommending qüestions based on previously failed qüestions. *Accuracy*, *Precision*, *Recall* and *Accuracy per class* metrics will be used by the creation of a *confussion matrix*. Also, the diferent obtained results will be commented as well as the diferent strategies that can be employed to improve the obtained results.

Index Terms— Machine Learning, Language processing, classification, histogram, Accuracy, Precision, Recall, Accuracy per class, Confusion Matrix, Unity



1 INTRODUCCIÓ

AQUEST document és una descripció del procés de creació d'un mòdul de processament de text per a ser emprat en la creació de videojocs amb l'aplicació d'intel·ligència artificial. Al llarg del document, s'aniran comentant les diferents parts clau del projecte, incloent els problemes trobats en la seva realització així com la seva resolució.

1.1 Motivació

Actualment a les universitats es pot trobar una part notable d'alumnes que mostren un baix interès en els estudis. Aquests alumnes mostren un baix rendiment acadèmic i, sovint, bé associat amb no prestar atenció a les lliçons o fins i tot en directament la no assistència a quan es tracta de lliçons presencials. Aquesta falta d'interès es produeix tant a dintre com a fora de les aules. Els alumnes amb falta d'interès per a les lliçons que hi assisteixen, sovint se'ls pot trobar utilitzant el seu telèfon intel·ligent

durant les classes.

En els últims anys, la societat ha mostrat un increment en la quantitat de persones que juguen a videojocs i gràcies a l'avanç tecnològic, qualsevol amb un telèfon intel·ligent pot disfrutar d'una àmplia varietat de videojocs, a la que la bona part dels jugadors hi inverteixen una quantitat de temps notòria per a jugar-hi.

Aquest treball proposa la creació d'un videojoc dirigit a persones que volen adquirir i reforçar els seus coneixements però troben difícil realitzar l'estudi independent o prestar atenció a les lliçons.

El plantejament original de l'objectiu del projecte consisteix en la realització d'un joc per al aprenentatge que, aplicant aprenentatge computacional, reforçés els coneixements del jugador. Aquest reforç, però, es centra en les àrees en les que els coneixements del jugador són més febles. Dit d'una altra manera: el jugador milloraria els seus coneixements degut a que si no respon correctament les preguntes plantejades de certa àrea, no pot avançar.

1.2 Fases del projecte

El projecte es divideix en tres fases. La primera fase

- E-mail de contacte: oriol.catalan@e-campus.uab.cat
- Menció realitzada: Computació
- Treball tutoritzat per: Jorge Bernal (ciències de la computació)
- Curs 2018/19

consisteix en la creació d'un mòdul que permeti processar la base de dades. La segona fase consisteix en l'ampliació del mòdul per a l'aplicació d'aprenentatge computacional en el llenguatge natural. La tercera, consisteix en la incorporació del mòdul a un projecte de joc en 2D.

1.3 Requeriments i objectius

Els requeriments del sistema es classifiquen en requeriments funcionals i requeriments no funcionals.

Els requeriments han estat agrupats segons la fase a la que pertanyen.

Taula 1 - Requeriments

Fase	ID	Requeriment	Tipus
1	R1	Mòdul per a processar la base de dades (llegir, escriure,...)	Funcional
2	R2	Mòdul per a la classificació i la recomanació de preguntes	Funcional
1, 2	R3	Creació d'un mòdul per a la realització de proves de verificació del comportament i rendiment del mòdul.	Funcional
3	R4	Interfície gràfica <i>user-friendly</i>	Funcional
1, 2	R5	Creació d'una base de dades SQL	No funcional
3	R6	Compilació del projecte per a múltiples plataformes	No funcional
3	R7	Creació d'un servidor per al control d'usuaris i de permisos.	No funcional

Per a la realització del projecte existeixen uns requeriments de *hardware* i *software*. Els requeriments del sistema es separen en requeriments mínims i requeriments recomanats.

Els requeriments mínims de sistema són:

- Sistema Operatiu: Windows 7, macOS, Linux o Android 7.0
- Memòria RAM: 2 GB
- Espai en disc: 150MB

Els requeriments recomanats de sistema són:

- Sistema Operatiu: Windows 7 o posterior, macOS, Linux o Android 7.0
- Memòria RAM: Mínim 4 GB
- Espai en disc: 300MB

Per a la realització del projecte s'han utilitzat:

- Ordinador portàtil:
 - Windows 10 64-bit
 - Intel Core i7-3537U
 - 8GB de RAM
 - SSD 500 GB
- Ordinador de sobretaula:
 - Windows 10 64-bit

- AMD Ryzen 5 1600
- 16GB de RAM
- SSD 750GB + 1TB HDD 7200rpm

Pel que fa als objectius, degut a la natura canviant dels projectes, durant el seu desenvolupament poden sorgir nous objectius o fins i tot, poden canviar els objectius originals. És molt important que els objectius del projecte estiguin ben definits tan aviat com sigui possible.

A continuació es mostren en forma de taula els objectius del projecte amb una estimació del seu grau de prioritat:

Taula 2. Objectius de projecte i prioritats.

Nº	Objectiu	Prioritat
1	Implementar un mòdul de processat de llenguatge i classificació	Alta
2	Disseny de la base de dades en SQL	Baixa
3	Creació de la interfície gràfica	Mitja
4	Disseny d'interfície fàcil i intuïtiva	Mitja
5	Implementació de la plataforma de proves	Alta
6	Implementació d'una xarxa neural per al processament del llenguatge	Baixa

1.3 Eines utilitzades

Per aquest projecte hi ha una gran diversitat d'eines a la nostra disposició per a les diferents parts.

1.3.1 Motor de videojoc

Per a la part d'interfície gràfica i de joc, teníem a la nostra disposició eines com Unreal Engine, Unity, i Godot entre altres. Donades les característiques de cada motor de videojocs, llenguatges de programació utilitzats i popularitat, es va decidir que per aquest projecte es la millor opció és Unity, ja que fa us de JavaScript i C#, amb els quals ja estic familiaritzat, a més del accés a una gran quantitat de llibreries i projectes d'exemple que eviten haver de desenvolupar la interfície gràfica des de zero donat que no és la part central del projecte. A més a més, Unity ofereix la possibilitat de compilació de projecte per a diverses plataformes.

1.3.2 Llenguatge de programació

Per a la part d'aprenentatge computacional es troben a la nostra disposició eines de programació com Python, Java, i Matlab. És molt senzill trobar exemples d'aprenentatge computacional en Python, i malgrat ser més senzill treballar en Python, no teníem clara la seva incorporació a un projecte que probablement utilitzi un llenguatge diferent. Per això, després de fer una recerca breu, es va decidir emprar C# degut a la seva similitud amb Java i per tant simplicitat a l'hora de convertir de

Java a C# estructures de codi per a us d'intel·ligència artificial i aprenentatge computacional a partir del exemples trobats tant a fòrums com a llibres i articles acadèmics.

1.3.3 Entorn de desenvolupament

Un cop escollits el motor de videojoc i el llenguatge a utilitzar, s'ha escollit el motor de desenvolupament (IDE). En el nostre cas, donat el motor de videojoc i el llenguatge escollits prèviament, es va decidir sense donar-hi masses voltes utilitzar Visual Studio 2017 Enterprise obtingut mitjançant el programari d'accés gratuït per a estudiants de la universitat autònoma amb col·laboració amb Microsoft. La raó per la que s'ha escollit és que Unity pot ser descarregat conjuntament amb el procés d'instal·lació de Visual Studio configurant automàticament Visual Studio per a us de Unity. A més, donat que C# és un llenguatge desenvolupat per Microsoft per a competir amb Java, Visual Studio esdevingué des de el nostre punt de vista la millor opció per a programar en C#.

1.3.4 Altres eines

Per al control de versions de codi i informes, així com portabilitat, es va decidir en un inici utilitzar GitHub que ja havíem fet servir amb anterioritat, però a meitat del projecte es va decidir canviar a Google Drive, degut a problemes sorgits durant el desenvolupament. També s'ha utilitzat Google Scholar per a la cerca d'informació i posterior cita de la mateixa a la bibliografia.

En els següents apartats s'exposaran: les eines emprades per a la realització del projecte; les metodologies escollides per a seguir en el desenvolupament del projecte; els problemes que han sorgit durant la realització, i les solucions emprades per a resoldre'ls; i els experiments realitzats i els resultats obtinguts.

2 ESTAT DE L'ART

Actualment ja existeixen eines que permeten aplicar el concepte de gamificació en l'aprenentatge, com ara:

- **Kahoot**[1] és una plataforma actualment bastant popular que permet la creació d'enquestes interactives amb participació del públic. S'accedeix a les enquestes mitjançant un codi des de qualsevol dispositiu. Quan s'han registrat les respostes, es mostrarà la que ha rebut més vots.
- **Socrative**[8] és una eina de suport a l'aprenentatge que permet la creació d'exercicis, quizz, i informes per a la monitorització del alumnes a les aules. Aquesta aplicació requereix que tant l'alumne com el professor es creïn un compte. Requereix d'adquisició de llicència.
- **PlayBrighter**[7] és una aplicació que requereix que el personal docent generi "missions" i objectius que hauran d'assolir els alumnes i permet que es proposin preguntes als alumnes.
- **Goose Chase**[10] és una aplicació que permet la crea-

ció de missions, i objectius a completar amb compatibilitat amb la geo-localització.

Aquestes aplicacions són serveis web d'educació que es comporten com jocs i requereixen d'un dispositiu amb connexió a internet per a utilitzar-les. Tot i així aquestes eines no permeten que l'estudiant, pugui posar a prova els seus coneixements sense la supervisió d'un membre de l'equip docent permetent que aquest pugui avançar de manera lúdica en la matèria..

Donat que avui dia un percentatge de gent significativa juga a videojocs, molts dels quals al mòbil durant els trajectes d'anada i tornada de les lliçons, el concepte base d'aquest treball pretén explotar aquest fet. Es proposa que, si els estudiants tenen a la seva disposició un joc que sigui prou interessant com a joc, i que a més a més estigui relacionat amb el seu camp d'estudi, ajudarà al estudiant que el jugui a consolidar i incrementar els seu coneixement.

3 METODOLOGIES ESCOLLIDES

En aquest apartat s'expliquen les metodologies i algorismes emprats, així com la justificació de perquè han estat escollides.

3.1 Algorismes

Per al mòdul d'intel·ligència artificial, donat que es tracta bàsicament d'un sistema recomanador, s'utilitzarà algorismes de clusterització per a classificar les preguntes per categories a més de algorismes de processament del llenguatge.

Es va realitzar una recerca sobre les diferents estratègies que es poden seguir per al processament de llenguatge. Finalment, per al processament de text es va escollir emprar el model de *Bag of Words* (BoW) donat que és un model que ja he estudiat i emprat anteriorment, i es poden trobar exemples de de com aplicar-lo per al processament del llenguatge així com exemples de recuperació d'informació. També, es va decidir que per a la classificació i recomanació de les preguntes segons la categoria s'utilitzarà la comparació d'histogrames generat a partir del model de BoW.

Per a entrenar i validar el model, es va dividir la base de dades amb les preguntes en dos parts. Una part va ser destinada a l'entrenament del model, mentre que l'altre va ser destinada a realitzar proves per a comprovar l'eficàcia. Com es veurà en l'apartat referent als experiments resultats, la mida d'aquests dos sets de dades es va variar en diferents proves, però la proporció de preguntes per a cada set es va mantenir uniforme per a cadascuna de les categories presents a les dades llegides per tal de mantenir un contingut equilibrat i així evitar que una de les categories rebi overfitting.

3.2 Metodologies de treball

Per a la realització del treball, es van dividir les fases en objectius a assolir. Aquests objectius ja han estat esmenats amb anterioritat en el document i petant no els tornarem a esmenar en aquest apartat, així que s'explicarà el funcionament general del projecte.

3.2.1 Metodologia del mòdul recomanador

El mòdul recomanador processarà la base de dades i la dividirà el conjunt de dades utilitzat en dos subconjunts: entrenament i test.

BoW, quan aplicat al processament del llenguatge, habitualment compta el nombre d'ocurrències de cada paraula del vocabulari (diccionari). En el nostre cas en particular, no contem les repeticions de les entrades del vocabulari en el moment de la seva creació, sinó que generem entrades de diccionari amb valor 0. Aquest diccionari s'utilitza per a generar els histogrames de les preguntes per a poder-ne realitzar comparacions més endavant. Així doncs, cada pregunta tindrà associada una còpia del diccionari (vocabulari).

Amb un segon processat de les preguntes del conjunt d'entrenament, s'actualitzen els valors a les entrades del diccionari associat a la pregunta segons el nombre d'ocurrències que tenen a la pregunta associada.

Un cop s'ha actualitzat els diccionaris de cada pregunta, es procedeix a calcular l'histograma mitjà de cada categoria. Aquest histograma mig consisteix en una còpia del diccionari inicial on les entrades obtenen el valor del mig de vegades que apareixen en les preguntes associades a cada categoria. Un cop realitzat el càlcul dels histogrames mig, considerem per acabada la fase d'entrenament i procedim amb la fase de proves.

Per a la fase de proves generem, amb el diccionari inicial obtingut durant la fase d'entrenament, els diccionaris "en blanc" de cadascuna de les preguntes del conjunt de test. Aquests diccionaris s'actualitzaran seguint el mateix procediment que en la fase d'entrenament, però en el cas que alguna paraula present al conjunt de test no es trobes present al conjunt d'entrenament, aquesta paraula nova no es tindrà en compte. Igual que en el cas del conjunt d'entrenament, els diccionaris de les preguntes serviran com a histogrames.

Quan ja hem generat els histogrames per a totes les preguntes i per a les categories, procedim a calcular la distància euclidiana entre els histogrames de les preguntes del conjunt de test i els histogrames de les categories. Aquesta primera comparació ens proporcionarà una predicció sobre la categoria a la que pot pertànyer la pregunta.

Un cop obtenim la predicció de la categoria a la que per-

tany la pregunta, procedim a calcular la distància euclidiana entre l'histograma d'una pregunta al atzar i els histogrames de les preguntes pertanyents a la categoria predita. La comparació d'histogrames entre preguntes es va decidir a limitar-la a la categoria predita per a obtenir una recomanació més ràpida que comparant amb el conjunt total de preguntes.

Finalment, es mostra la pregunta que sigui més similar (tingui la menor distància euclidiana entre els histogrames) a la pregunta escollida.

3.2.2 Metodologia del la interfície gràfica

Per a interactuar amb els usuaris es va decidir implementar una interfície gràfica, i donat que la idea original consistia en la creació d'un videojoc, es va escollir fer us de Unity i es va descarregar una plantilla de demostració d'ús lliure que es pot modificar i se li poden afegir mòduls.

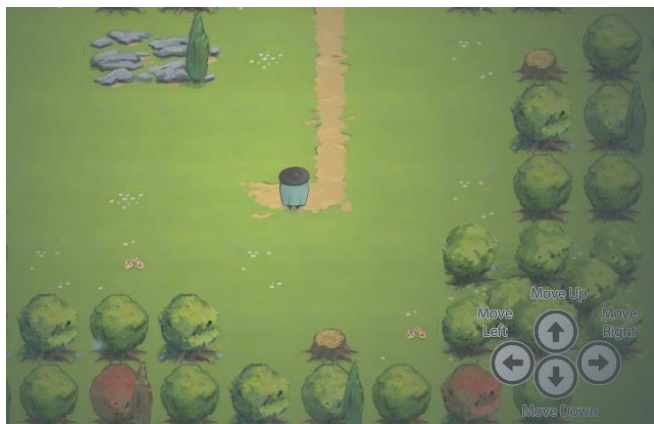
En aquest mòdul de joc, el jugador controla un personatge que es pot desplaçar lliurement pel mapa sempre que no col·lideixi amb objectes invariants del mapa (arbres, pedres,...). El jugador es desplaçarà cap a uns objectes lluminosos que, al entrar en contacte amb el jugador, realitzaran una pregunta que el jugador haurà de respondre. Si el jugador respon correctament, l'objecte lluminós serà destruït, mentre que si s'equivoca, l'objecte realitzarà una nova pregunta similar a la que el jugador ha fallat. El jugador haurà de destruir tots els objectes lluminosos per a passar al següent nivell i poder continuar amb el joc.

Es planteja, a més, que aleatòriament aparegui un personatge no jugable (NPC) que realitzi preguntes al jugador, i si aquest les respon correctament, el jugador obtingui algun benefici, com ara recuperació dels punts de salut, un escut per a guarir-se del mal generat al fallar la resposta a la pregunta, o recuperació completa dels punts d'energia que siguin necessaris per a continuar avançant.

També es vol implementar un sistema de punts de salut i energia, de manera que per a les preguntes fallides, es produeixi una pèrdua de punts de salut, podent provocar la mort del personatge. Això evitarà que un jugador pugui respondre aleatòriament sense repercussions.

Les preguntes es planteja que incrementin la dificultat a mida que s'avança el joc, pel que es pretén incloure un sistema de punts d'experiència que permetin a un jugador respondre a preguntes adequades al seu nivell. Així, s'evitarà que un jugador perdi tots els punts intentant respondre preguntes massa avançades.

El mòdul de joc és part d'un joc basat en fitxes. Això vol dir que tant el mapa del joc com els seus elements es componen per "fitxes" que indiquen que es cada cosa. A continuació es mostren imatges del joc:



Il·lustració 1 - Demo de joc mode dia



Il·lustració 2 - Demo de joc mode nit

Donat que el procés de creació de model per recomanar preguntes consumeix temps i recursos que fa que no sigui adequat realitzar la creació d'un nou model cada vegada, es realitzarà una vegada la creació del model, i serà desada en un document apart que es llegirà al llençar el joc conjuntament amb la base de dades de preguntes amb els seus histogrames per a escurçar el temps d'execució i així no entorpir l'experiència de joc.

El joc és simple de jugar: el jugador només s'haurà de desplaçar pel mapa per poder avançar, trobar els objectes lluminosos, i respondre preguntes per a pujar de nivell.

Per a comprovar que el mòdul recomanador funciona i obtenir resultats de l'eficiència del l'estratègia aplicada, es va desenvolupar un projecte a part per a poder mostrar els resultats per consola. Aquesta aplicació de consola mostra la matriu de confusió de la prova, amb diferents mètriques per a validar el model de classificació segons la categoria de la pregunta. També es mostra un exemple d'una pregunta aleatòria i, basant-se en la categoria a la que ha estat assignada, una pregunta similar. A més, també servirà per a escriure en documents externs els histogrames mitjos de les categories, així com les preguntes i els seus histogrames corresponents. Tot seguit es mostren exemples de l'aplicació per consola:

```
Sports & Leisure:
Training: 882 Test: 378

Geography:
Training: 904 Test: 387

History & Holidays:
Training: 856 Test: 367

Science & Nature:
Training: 864 Test: 370

Music:
Training: 852 Test: 365

General:
Training: 886 Test: 380

Confusion Matrix
108 37 31 36 70 37
21 233 48 68 5 85
38 21 114 19 32 75
33 22 30 103 13 50
128 23 81 81 219 15
50 51 63 63 26 118
```

Il·lustració 3 - Aplicació de consola: Distribució de les preguntes i matriu de confusió

```
Accuracy: 39,9465597862391 %

Sports & Leisure
Recall: 30,1587301587302 %
Precision: 32,618025751073 %
Class Accuracy: 64,2320085929108 %

Geography
Recall: 52,7131782945737 %
Precision: 50,7462686567164 %
Class Accuracy: 70,18779342723 %

History & Holidays
Recall: 29,9180327868852 %
Precision: 36,8686868686869 %
Class Accuracy: 66,8903803131991 %

Science & Nature
Recall: 33,6032388663968 %
Precision: 45,3551912568306 %
Class Accuracy: 69,3735498839907 %

Music
Recall: 64,6090534979424 %
Precision: 41,8666666666667 %
Class Accuracy: 66,2971175166297 %

General
Recall: 28,8537549407115 %
Precision: 30,4166666666667 %
Class Accuracy: 63,2804232804233 %

What song links Little Eva and Kylie Minogue
What are these: Ceres, Juno, Iris, and Flora?
```

Il·lustració 4 - Aplicació de consola: mètriques, pregunta aleatòria, i pregunta recomanada

Amb el mòdul de consola creat, es va procedir a realitzar proves i experiments per a determinar el rendiment del model en diversos escenaris.

4 EXPERIMENTACIÓ I RESULTATS

En aquesta secció, s'exposen les propostes d'experimentació realitzades per a provar i confirmar el funcionament del mòdul recomanador. Per a la realització de les proves, donat que es requerien dades amb una etiqueta s'ha fet servir una base de dades de preguntes

del joc Trivial. Donat que la quantitat de preguntes no era equilibrada a totes les categories, es va decidir limitar les categories que s'utilitzarien. També es va manipular seleccionar una quantitat de preguntes similar per a cada categoria ja que, per exemple, la categoria *General* posseeix un volum de dades major que qualsevol altre categoria.

Taula 3 Nom, ID assignada, i nombre de preguntes per categoria

CATEGORIA	ID	# PREGUNTES
SPORTS & LEISURE	1	1260
GEOGRAPHY	2	1291
HISTORY & HOLLIDAYS	3	1223
SCIENCE & NATURE	4	1235
MUSIC	5	1217
GENERAL	6	1266

Per a mesurar l'eficàcia del model s'ha generat una matriu de confusió a partir de la que es calculen les mètriques per a cada categoria. Les mètriques utilitzades són: *Precision*, *Recall*, i *Accuracy*. La mètrica *Precision* ens serveix per a mesurar les vegades que una predicció ha estat correcte; *Recall* mesura les vegades que s'ha identificat correctament una categoria; *Accuracy* mesura les prediccions que el nostre model ha encertat. A més a més, s'ha calculat l'*Accuracy* global del model per a cadascun dels diferents escenaris.

4.1 Experiments realitzats

Com s'ha mencionat anteriorment, s'han realitzat diferents experiments sobre el model per a comprovar-ne el funcionament. Aquests experiments inclouen la variació en la proporció de dades destinades a entrenament i test, estratègies per a eliminar "soroll" en el diccionari provenint de les preguntes processades, i modificació de la quantitat de categories utilitzades. Cada prova ha estat executada diverses vegades i s'ha escollit mostrar els resultats que no han obtingut el millor ni el pitjor rendiment.

Per a tots els experiments, s'han filtrat els caràcters no alfanumèrics en la generació dels histogrames, ja que no aporten informació sobre les categories.

4.1.1 Cas base

Per a aquesta prova es va considerar que totes les paraules del conjunt d'entrenament apareixerien al diccionari, sense omissions ni restriccions.

Aquesta prova serveix com a cas base de referència. A partir d'aquest cas d'experimentació, s'han aplicat modificacions en el tractament de les dades que donen lloc als altres casos d'experimentació.

En aquesta prova, es va decidir utilitzar un 80% de les dades com a entrenament i un 20% de les dades com a test. Un cop finalitzada la seva execució, es va obtenir la següent matriu de confusió:

Taula 4 Matriu de confusió (cas base)

	1	2	3	4	5	6
1	70	13	29	15	35	25
2	18	158	22	47	4	67
3	23	10	78	12	17	43
4	18	15	8	71	8	42
5	82	23	48	54	164	8
6	41	39	59	48	15	68

Observant la matriu de confusió del experiment, podem veure que les categories 2 i 5 s'han classificat millor que les categories 1, 3, 4, i 6. Per a comprovar l'eficàcia general del model, es calcula l'*Accuracy* general.

$$Accuracy = 40,68\%$$

Així doncs, observem que el model erra la predicció de les categories de les preguntes aproximadament un 60% de les vegades.

Seguidament, procedim a calcular les mètriques *Precision*, *Recall*, i l'*Accuracy* de cada categoria, que ens permetrà veure el rendiment del model a l'hora de classificar cadascuna de les categories.

Taula 5 Mesures obtingudes (Cas base)

	PRECISION	RECALL	ACCURACY
1	27,78%	37,43%	67,07%
2	61,24%	50,00%	70,24%
3	31,97%	42,62%	69,20%
4	28,74%	43,83%	69,52%
5	67,49%	43,27%	67,44%
6	26,88%	25,19%	61,14%

Un cop calculades les mètriques, observem que la categoria que assoleix un millor rendiment és la categoria 2, mentre que la categoria que obté un pitjor rendiment és la categoria 6.

Per últim, es calcula la distància mitja dels histogrames de les preguntes del conjunt de test als histogrames mitjos de cada categoria en segons la categoria a la que pertanyen (cal tenir en compte que la màxima distància és 1 i la mínima es 0):

Taula 6 Distància mitjana de a les categories

	1	2	3	4	5	6
1	0,00077	0,00062	0,00081	0,00080	0,00086	0,00067
2	0,00087	0,00051	0,00085	0,00080	0,00100	0,00065
3	0,00080	0,00058	0,00078	0,00079	0,00090	0,00064
4	0,00080	0,00056	0,00081	0,00076	0,00091	0,00064
5	0,00080	0,00069	0,00085	0,00084	0,00084	0,00073
6	0,00083	0,00055	0,00079	0,00078	0,00095	0,00062

Així dons, amb els resultats del cas base obtinguts, procedim es va procedir a l'execució dels altres experiments

4.1.2 Sense preposicions

Les preposicions són paraules invariables que serveixen per a establir una relació de dependència entre dues o més paraules. Encara que el significat d'una frase varia en funció de la preposició utilitzada, no és característica de cap categoria en particular, i petant, el valor que poden aportar a l'hora de classificar preguntes pel seu contingut, no és rellevant. Per aquest motiu, en aquest experiment es va decidir obviar les preposicions.

L'execució d'aquesta prova va produir la següent matriu de confusió:

Taula 7 Matriu de confusió (Sense preposicions)

	1	2	3	4	5	6
1	81	21	32	23	39	15
2	22	137	29	40	4	56
3	30	21	68	11	18	44
4	16	18	24	77	8	38
5	73	23	44	51	161	6
6	30	38	47	45	13	94

Igual que en el cas base, es va observa que les categories que s'han classificat millor segons la matriu de confusió, són les categories 2 i 5, indicant així que malgrat el canvi en el diccionari, ambdues categories poden tenir vocabulari més característic que la resta. Per altre banda, es va observar que les categories 2 i 5 obtenien pitjors resultats que anteriorment, i la resta de categories millorava els seus resultats. Per a poder comparar més còmodament amb el cas base, també se'n va calcular l'Accuracy general del model.

$$Accuracy = 41,28\%$$

En aquest cas d'experimentació, es va observar una millora en l'Accuracy respecte el cas base. Cal destacar, però que en les diferents execucions del mateix escenari de proves, es van obtenir resultats diferents que feien que l'Accuracy general oscil·lés entre el 39% i el 43%.

Taula 8 Mesures obtingudes (Sense preposicions)

	PRECISION	RECALL	ACCURACY
1	32,14%	38,39%	67,25%
2	53,10%	47,57%	69,44%
3	27,87%	35,42%	67,32%
4	31,17%	42,54%	69,28%
5	66,26%	44,97%	68,90%
6	37,15%	35,21%	65,05%

En aquest experiment, es va comprovar que mentre les categories 2 i 3 resulten perjudicades per la decisió d'ignorar les preposicions, la resta de categories se'n beneficien, aconseguint així que no hi hagi la disparitat entre el rendiment de les categories que existeix al cas base, el que ha derivat en la millora del rendiment general del model.

4.1.3 Sense "the"

La paraula "the" del diccionari anglès és un determinant que referencia a una o més coses ja mencionades o de coneixement popular, o també a una per apuntar a una clàusula o frase de qualificació o definició. En aquesta prova, es va decidir excloure la paraula "the" donat que BoW no té en compte l'ordre de les paraules, i petant "the" ens suposa una paraula que es pot repetir al llarg de les categories i no ens aporta cap informació rellevant.

De l'execució de la prova, en vam poder obtenir la següent matriu de confusió:

Taula 9 Matriu de confusió ("the")

	1	2	3	4	5	6
1	70	22	19	18	26	14
2	25	152	15	52	3	59
3	37	8	100	13	30	40
4	10	11	7	62	7	12
5	67	18	50	41	150	9
6	43	47	53	61	27	119

Novament, vam poder observar en aquest cas que la categoria 2 es veia perjudicada per la decisió escollida, però a diferència del cas sense preposicions, la categoria 4 també es va veure perjudicada. En aquest cas, les categories 1, 3, 5, i 6 es van veure beneficiades pel canvi. De nou, es va calcular-ne l'Accuracy general del experiment.

$$Accuracy = 43,62\%$$

Amb aquest resultat es pot afirmar que per al model, des de la perspectiva general del mateix, resultava beneficiat d'aquest canvi. Per a les diferents execucions d'aquest escenari, els resultats obtinguts resultaven en una oscil·lació de l'Accuracy general entre el 40% i el 45%. Per a comprovar com afecta el canvi a les diferents categories,

es va procedir a calcular les mètriques per a cadascuna:

Taula 10 Mesures obtingudes (sense preposicions)

	PRECISION	RECALL	ACCURACY
1	27,78%	41,42%	69,91%
2	58,91%	49,67%	71,52%
3	40,98%	43,86%	70,59%
4	25,10%	56,88%	73,79%
5	61,73%	44,78%	70,14%
6	47,04%	34,00%	64,15%

Si observem *Precision* i *Recall*, podem veure com en alguns casos, ambdues milloren, significat una millora en la validesa de les prediccions de categoria realitzades, i una millora en la identificació de les categories..

4.1.2 Variació en la distribució de les dades

Com a últim experiment, es va escollir comprovar com afecta la divisió de les dades en *training* i *test* al rendiment del model. En aquest cas, es vol comprovar el rendiment que s'obté quan les dades es reparteixen en un rati de 50:50, un rati de 70:30, i un rati de 90:10 per a *training* i *test* respectivament.

Per a cap dels següents experiment, no es va aplicar cap altre alteració.

50% training, 50% test

Per aquest cas, es va repartir les dades a parts iguals entre *training* i *test*. La seva execució va produir els següents resultats.

Taula 11 Matriu de confusió (50:50)

	1	2	3	4	5	6
1	186	63	70	61	104	68
2	47	393	76	126	15	141
3	94	26	207	31	53	130
4	28	32	35	175	21	62
5	191	51	143	112	391	23
6	85	81	81	111	25	209

A partir de la matriu de confusió obtinguda, es calculen novament l'*Accuracy* general del model en aquest escenari i les mètriques per a mesurar el rendiment per a cada categoria:

$$Accuracy = 41,66\%$$

Per a comprovar com afecta el canvi a cada categoria, es va procedir a calcular les mètriques de validació:

Taula 12 Mesures obtingudes (50:50)

	PRECISION	RECALL	ACCURACY
1	29,48%	33,70%	65,81%
2	60,84%	49,25%	70,35%
3	33,82%	38,26%	67,87%
4	28,41%	49,58%	71,61%
5	64,20%	42,92%	67,90%
6	33,02%	35,30%	65,92%

Amb les mètriques calculades, es va observar com el rendiment millora en la majoria de categories.

70% training, 30% test

Per aquest cas, es va repartir les dades en un 70% per a *training* i un 30% per a *test*. La seva execució va produir els següents resultats.

Taula 13 Matriu de confusió (70:30)

	1	2	3	4	5	6
1	35	12	10	13	16	8
2	11	75	13	22	1	31
3	19	6	44	4	9	33
4	10	6	11	41	10	10
5	36	9	29	20	76	1
6	15	21	15	23	9	43

A partir de la matriu de confusió obtinguda es van calcular l'*Accuracy* general en aquest escenari:

$$Accuracy = 42,03\%$$

i les mètriques per a mesurar el rendiment per a cada categoria:

Taula 14 Mesures obtingudes (70:30)

	PRECISION	RECALL	ACCURACY
1	27,78%	37,23%	67,67%
2	58,14%	49,02%	70,40%
3	36,07%	38,26%	67,82%
4	33,33%	46,59%	70,88%
5	62,81%	44,44%	69,16%
6	34,13%	34,13%	65,42%

Amb les mètriques calculades, es va observar com el rendiment es manté en una franja de valors molt similar a la del cas anterior i la del cas base, donant fonament a la idea de l'excés vocabulari sense importància (soroll) dins el conjunt de dades.

90% training, 10% test

En l'últim cas que es va provar, es van repartir les dades en un 90% per a *training* i un 10% per a *test*. La seva execució va produir els següents resultats.

Taula 15 Matriu de confusió (90:10)

	1	2	3	4	5	6
1	40	9	17	11	15	11
2	8	78	13	27	1	40
3	12	10	35	6	8	16
4	10	4	8	37	5	12
5	41	12	30	32	86	3
6	15	16	19	10	6	44

A partir de la matriu de confusió obtinguda, i com en els experiments anteriors es van calcular l'*accuracy* general en aquest escenari:

$$Accuracy = 42,84\%$$

i les mètriques de rendiment per categoria:

Taula 16 Mesures obtingudes (90:10)

	PRECISION	RECALL	ACCURACY
1	31,75%	38,83%	68,23%
2	60,47%	46,71%	69,57%
3	28,69%	40,23%	69,72%
4	30,08%	48,68%	71,91%
5	71,07%	42,16%	67,65%
6	34,92%	40,00%	68,38%

En aquest cas, es va apreciar un cop més que les mètriques es mantenen en el mateix rang de valors.

Per aquest últim experiment, s'ha deduït que donat el volum de dades utilitzat, i la quantitat de paraules sense pes per al vocabulari de cada categoria, les variacions en la distribució de les dades en *training* i *test* no suposen una gran diferència entre 50:30, 70:30, 80:20, i 90:10.

5 CONCLUSIÓ

5.1 Problemes i solucions

En aquest apartat, s'exposaran els problemes trobats durant el desenvolupament del projecte així com les solucions aplicades per a resoldre'ls.

5.1.1 Problemes trobats

Durant el procés de desenvolupament del projecte es van anar trobant certs problemes.

El primer problema va sorgir va ser l'obtenció de dades

per a entrenar l'algorisme. Donat que les preguntes que podrien proporcionar els membres de l'equip docent resultarien en estar molt limitades en nombre i petant resultarien poc apropiades.

El segon problema que va sorgir va ser el plantejament de com procedir a comparar els histogrames. Donat que la comparació pot ser realitzada de moltes formes diferents es va dubtar de quina resultaria més efectiva.

El tercer problema, va ser inesperadament, la creació de l'aplicació de consola.

El quart problema van ser la creació de la matriu de confusió així com el càlcul de les mètriques necessàries per a l'avaluació del mòdul.

5.1.2 Solucions aplicades

Per a cada problema que va aparèixer, es van buscar diferents solucions.

Per al problema de les dades a tractar, després de pensar en com aconseguir prou dades, o com substituir-les per unes que pugin resultar equivalents, es va decidir cercar una base de dades de preguntes de Trivial. Cadascuna d'aquestes dades conté categoria, pregunta, i resposta. Aquest canvi, va resultar en que es desistís de la creació d'una base de dades pròpia, ja que el volum de dades a introduir requeriria massa temps.

Per al problema referent a la comparació dels histogrames, després de la reunió pertinent amb el tutor es va decidir emprar la distància euclidiana ja que resultava simple i prou eficaç.

Per al tercer problema, després d'una investigació per internet de perquè no es podia mostrar per consola els resultats de les proves des de el mòdul creat, es va crear un nou projecte que sí que permet la utilització de la consola per a mostrar els resultats i comprovar que el mòdul funciona correctament.

5.2 Demo de joc

La demo del joc, es troba encara ara en fase de desenvolupament. La demo permet al jugador desplaçar-se en qualsevol direcció del mapa, sempre que no col·lideixi amb algun objecte (arbre, roca,...). També té activades unes caselles que segons el moment en el que el jugador hi faci contacte, el personatge morirà i es tornarà a començar. Pel que fa als objectes lluminosos que han de realitzar les preguntes, s'ha generat l'event de destrucció d'aquests en quant s'hi entra en contacte.

Es pot trobar una demostració del joc al enllaç de la referència [20]

5.3 Conclusions i futur

Aquest treball m'ha permès aprofundir i refrescar els meus coneixements d'aprenentatge computacional i xarxes neurals, adquirits al llarg del grau, així com adquirir experiència en l'àmbit de desenvolupament d'aplicacions amb interfície gràfica..

El projecte encara es troba en una fase embrional i no pot ser portat encara cara al públic per a que hi juguïn. No només la interfície de joc es troba poc desenvolupada, sinó que el mòdul recomanador de preguntes, encara ha de millorar per a que realment sigui efectiu per a ajudar a adquirir coneixements.

Durant el procés de desenvolupament del projecte, es va plantejar l'ús de xarxes neuronals per al tractament de les preguntes, la seva classificació i la seva recomanació. Seguint aquesta idea, es va començar a desenvolupar un conjunt de classes que facilitessin la creació i ús de xarxes neuronals en C#, però degut a falta de temps es va aturar el desenvolupament d'aquesta estructura, i es va decidir a centrar els esforços en dur a terme la tasca original.

Finalment, ha quedat pendent el tractament de vocabulari desconegut al processar una pregunta nova. Es va plantejar que, quan una paraula desconeguda aparegui en una nova pregunta, incloure aquesta al diccionari del mòdul, tenint en compte a quina categoria s'ha predit que pertany la pregunta, ampliant el vocabulari del model.

AGRAÏMENTS

Voldria expressar el meu més sincer agraïment al meu tutor, Jorge Bernal del Nozal, per la seva paciència i comprensió, així com per la seva indispensable ajuda en la presa de decisions respecte les estratègies a seguir; a la meua família, que m'han deixat en treballar tranquil quan ho he necessitat; als meus amics, que m'han donat tot el seu suport i s'han ofert ajudar-me (en especial a l'Ana, que m'ha ajudat a redactar correctament els informes); i a Internet, per posar al meu abast una font inesgotable d'informació, tot i que a vegades ha estat contraproductiu per a la meua concentració.

BIBLIOGRAFIA

- [1] Kahoot! | Learning Games | Make Learning Awesome!. (2018). Retrieved from <https://kahoot.com/welcomeback/>
- [2] Top 3 alternatives a Kahoot. (2018). Retrieved from <http://el-proyector-de-clase.blogspot.com/2017/08/3-alternativas-kahoot.html>
- [3] BONACCORSO, Giuseppe. Machine learning algorithms. Packt Publishing Ltd, 2017.
- [4] SOARES, Fábio M.; SOUZA, Alan MF. Neural network programming with Java. Packt Publishing Ltd, 2017.
- [5] Evaluación de precisión del modelo. Retrieved from Amazon Machine Learning Guía del desarrollador Version Latest https://docs.aws.amazon.com/es_es/machine-learning/latest/dg/machinelearning-dg.pdf#evaluating-model-accuracy
- [6] Machine Learning Crash Course. Retrieved from Google Developers <https://developers.google.com/machine-learning/crash-course/>
- [7] Play Brighter. 2019. [Mobile application software]. Retrieved from

- <http://playbrighter.com/>
- [8] Coca, D. M., & Sliško, J. (2017). Software Socrative and smartphones as tools for implementation of basic processes of active physics learning in classroom: An initial feasibility study with prospective teachers. *European Journal of Physics Education*, 4(2), 17-24.
- [9] Schwaber, K., & Beedle, M. (2002). *Agile software development with Scrum* (Vol. 1). Upper Saddle River: Prentice Hall.
- [10] Goose Chase. 2019. [Mobile application software] Retrieved from <https://www.goosechase.com/>
- [11] Unity Tutorials. 2019. <https://unity3d.com/es/learn/tutorials>
- [12] ZHANG, Yin; JIN, Rong; ZHOU, Zhi-Hua. Understanding bag-of-words model: a statistical framework. *International Journal of Machine Learning and Cybernetics*, 2010, 1.1-4: 43-52.
- [13] JOULIN, Armand, et al. Bag of tricks for efficient text classification. arXiv preprint arXiv:1607.01759, 2016.
- [14] SEBASTIANI, Fabrizio. Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 2002, 34.1: 1-47.
- [15] IKONOMAKIS, M.; KOTSIANTIS, Sotiris; TAMPAKAS, V. Text classification using machine learning techniques. *WSEAS transactions on computers*, 2005, 4.8: 966-974.
- [16] RYTTER, M. Crochemore W.; CROCHEMORE, M. *Text Algorithms*. Oxford University Press, 1994.
- [17] GOLDSTONE, Will. *Unity game development essentials*. Packt Publishing Ltd, 2009.
- [18] OAK, Ji Won; BAE, Jae Hwan. Development of Smart Multiplatform Game App Using Unity 3D Engine for CPR Education'. *International Journal of Multimedia and Ubiquitous Engineering*, 2014, 9.7: 263-268.
- [19] MURRAY, Jeff W. *C# game programming cookbook for Unity 3D*. AK Peters/CRC Press, 2014.
- [20] CATALAN YARZA, Oriol. Demo de joc <https://drive.google.com/file/d/1cj6kRfvRmiSLtNfONtZxqaNbx2oai/view?usp=sharing>
- [21] HYPER LUMINAL GAMES. 2D Fantasy Forest Tileset. 2016 <https://assetstore.unity.com/packages/2d/environments/2d-fantasy-forest-tileset-19553>