

GRAU EN ENGINYERIA INFORMÀTICA

ESPECIALITAT EN COMPUTACIÓ

TREBALL DE FI DE GRAU

Construcció d'un dataset per a l'extracció de models de processos de negoci a partir de textos

GUILLEM PLA BERTRAN

Director
JOSEP CARMONA VARGAS

Codirector
LLUÍS PADRÓ CIRERA

Tutora del GEP
OLGA PONS PEREGORT

01/03/2022



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Facultat d'Informàtica de Barcelona



Resum

La definició formal de processos de negoci en estàndards com BPMN 2.0 és un pas crucial per a la transformació digital de les organitzacions. Moltes organitzacions tenen els seus processos documentats en text, i crear-ne els models formals té un elevat cost de personal especialitzat.

L'objectiu del projecte és crear un conjunt de dades anotades que relacionin textos amb els models formals que els descriuen, per tal que puguin ser usades per entrenar sistemes neuronals de Deep Learning per fer aquesta tasca.

S'usaran tècniques de *data augmentation* per crear textos sintètics a partir de models existents, i es realitzarà una prova de concepte amb un model neuronal simple sobre les dades obtingudes.

Índex

1	Introducció	5
2	Context	6
2.1	Problema a resoldre	6
2.2	Actors implicats	6
2.3	Business Process Model and Notation	7
2.3.1	Elements de BPMN	7
3	Justificació	11
4	Abast	13
4.1	Objectius	13
4.2	Possibles obstacles	14
5	Metodologia	15
5.1	Mètode Àgil	15
5.2	Eines	16
5.3	Validació	16
6	Planificació temporal	17
6.1	Definició de tasques	17
6.1.1	Recursos necessaris	17
6.1.2	Gestió del Projecte - GP	18
6.1.3	Generació del dataset - GD	19
6.1.4	Crear model Deep Learning - MD	20
6.1.5	Documentació - DO	21
6.2	Estimacions i Gantt	23
6.3	Plans alternatius i obstacles	25
7	Pressupost	27
7.1	Costos de personal	27
7.2	Costos per tasca	27
7.3	Costos genèrics	28
7.3.1	Maquinari	28

7.3.2	Programari	29
7.3.3	Teletreball	29
7.4	Total dels costos per personal i genèrics	30
7.4.1	Contingència	30
7.4.2	Imprevistos	30
7.5	Cost total del projecte	31
7.6	Control de gestió	31
8	Informe de sostenibilitat	32
8.1	Dimensió econòmica	32
8.2	Dimensió ambiental	33
8.3	Dimensió social	33
9	Solució plantejada	34
9.1	Eines utilitzades	34
9.2	Implementació	34
9.3	Solucions prèvies	34
9.4	Alternatives	34
9.5	Obstacles	34
10	Conclusions	35
11	Agraïments	36
	Bibliografia	37

Índex de figures

2.1	Representació dels diferents tipus d'Esdeveniments. Font: Elaboració pròpia.	8
2.2	Representació dels diferents tipus d'Activitats. Font: Elaboració pròpia. .	9
2.3	Representació dels diferents tipus de <i>Gateways</i> . Font: Elaboració pròpia. .	9
2.4	Representació dels diferents tipus de Connectors. Font: Elaboració pròpia.	10
4.1	Subobjectius de cada part del projecte. Font: Elaboració pròpia.	13
5.1	Metodologia Agile. Font: csgsolutions.	15
6.1	Primera part del diagrama de Gantt del projecte. Font: Elaboració pròpia.	24
6.2	Segona part del diagrama de Gantt del projecte. Font: Elaboració pròpia.	25

Índex de taules

6.1	Taula de les tasques. Font: Elaboració pròpia.	22
7.1	Rols del projecte i la facturació per hora. Font: Elaboració pròpia.	27
7.2	Taula del cost per cada tasca. Font: Elaboració pròpia.	28
7.3	Costos del maquinari. Font: Elaboració pròpia.	29
7.4	Costos pel programari. Font: Elaboració pròpia.	30
7.5	Costos de personal i genèrics. Font: Elaboració pròpia.	30
7.6	Taula amb els costos dels imprevistos. Font: Elaboració pròpia.	31
7.7	Taula amb els costos totals del projecte. Font: Elaboració pròpia.	31

1

Introducció

El diccionari de l'Institut d'Estudis Catalans [1] defineix un procés com una manera de descabdellar-se una acció progressiva. És a dir, que qualsevol conjunt d'accions que realitzem en el nostre dia a dia es pot definir com a un procés. De la mateixa manera, en una empresa es duen a terme una llarga llista d'accions que permeten aconseguir els objectius marcats.

Evidentment, en una empresa els processos no són tan simples com els que realitzem nosaltres diàriament. En conseqüència, és molt important que cada una de les tasques estigui ben definida.

La definició formal de processos de negoci en estàndards com BPMN 2.0 és un pas crucial per a la transformació digital de les organitzacions, ja que permet gestionar els processos que es duen a terme en una empresa. Moltes organitzacions tenen els seus processos documentats en text, i crear-ne els models formals té un elevat cost de personal especialitzat.

2

Context

Aquest treball de fi de grau (TFG) del Grau en Enginyeria Informàtica (GEI) de l'especialitat de Computació es fa amb modalitat A (centre), i es realitza en la Facultat d'Informàtica de Barcelona (FIB). El projecte pretén aprofundir en la recerca per aconseguir una millora en la creació de models BPMN a partir de textos en llenguatge natural.

2.1 Problema a resoldre

Com s'ha comentat anteriorment, el modelatge dels processos empresarials és una part vital a l'hora d'establir el funcionament d'un projecte. Moltes empreses tenen textos que expliquen com estan definits les accions a realitzar, però no ho tenen modelat en un llenguatge formal. El modelatge pot arribar a comportar fins al 60% del temps gastat en els projectes de gestió de processos. Per aquest motiu, en aquest treball s'intentarà proporcionar un programa que converteixi la documentació existent dels processos en diagrames BPMN.

2.2 Actors implicats

Els actors implicats en aquest projecte són el personal del projecte, és a dir, els directors Josep Carmona Vargas i Lluís Padró Cirera, la tutora del GEP Olga Pons Peregort i l'autor Guillem Pla Bertran.

També hi estan implicades les persones a les quals va dirigit el projecte. Aquestes són totes les persones que treballen amb BPMN en el món empresarial, principalment *Business Analysts*. Segons la Northeastern University [2] aquestes persones són els responsables d'entendre què es fa en el negoci, quins passos cal seguir per a completar els objectius, evaluar els processos per eficiència, cost i resultat, desenvolupar plans del projecte, entre altres coses.

Com que els *Business Analysts* són els que més ús en poden donar també són qui es beneficiarà d'aquest projecte. Ja que veuran com la seva feina és més senzilla de realitzar. I evidentment, també se'n beneficiaran les empreses que emprin aquesta tecnologia, ja que experimentaran un increment en la productivitat i una reducció dels recursos destinats al modelatge de processos.

2.3 Business Process Model and Notation

BPMN vol dir Business Process Model and Notation, en català Model i Notació de Processos Empresariums, és un estàndard per a la modelització de processos empresariums[3]. Aquest proporciona una notació gràfica per descriure un procés en un Business Process Diagram (BPD). Està basat en una tècnica de diagrama de flux molt similar als diagrames d'activitats de l'Unified Modeling Language (UML).

L'objectiu de BPMN és donar suport a la gestió de processos empresariums, tant per als usuaris tècnics com per als usuaris empresariums, per això proporciona una notació intuïtiva, però que alhora és capaç de representar una semàntica de processos complexa.

Recentment, la quantitat d'empreses que han començat a utilitzar BPMN ha crescut molt. Segons una enquesta del 2018 [4] el 64% dels negocis estan interessats en utilitzar aquest mètode per a simplificar els seus processos empresariums per a poder estalviar diners i millorar la productivitat.

A part de millorar l'eficiència hi ha altres motius que expliquen el perquè és necessari [5]. Augmenten la satisfacció dels clients, milloren la capacitat de resposta organitzativa i ajuden a complir amb noves regulacions, entre d'altres.

2.3.1 Elements de BPMN

Els models BPMN fan servir diagrames simples construïts a partir d'un conjunt d'elements gràfics. Els quatre elements bàsics són els *Events*, les *Activities*, les *Gateways* i els *Connectors*.

Esdeveniments

Un *Event* o Esdeveniment es representa amb un cercle i significa que alguna cosa passa durant el transcurs d'un procés. Poden tenir símbols que determinin el tipus d'esdeveniment que són. Principalment, es fan servir tres tipus d'esdeveniments:

- **Start event** o Esdeveniment inicial: És el que inicia tot el procés. Tots els processos comencen amb un esdeveniment d'aquest tipus.
- **Intermediate event** o Esdeveniment intermedi: Representa que alguna cosa passa entre l'inici i el final del procés.

- **End event** o Esdeveniment final: Representa el resultat final del procés. Tot procés té com a mínim un esdeveniment final.



Figura 2.1: Representació dels diferents tipus d'Esdeveniments. Font: Elaboració pròpia.

Activitats

Tal com explica Visual Paradigm [6], una *Activity* o Activitat és un "treball" que una empresa duu a terme en un procés de negoci. Les activitats poden ser atòmiques (Tasques) o descomponibles (Subprocessos). N'hi ha tres tipus bàsics:

- **Task** o Tasca: És una activitat atòmica dins d'un flux de processos. Es creen quan l'activitat no es pot desglossar a un nivell més detallat. En general, una persona o aplicacions duren a terme la tasca quan s'executi.
- **Sub-Process** o Subprocés: És una activitat composta que representa una col·lecció d'altres tasques i subprocessos. Per facilitar les comunicacions, no es vol que un diagrama de processos empresarials sigui complex. Mitjançant l'ús de subprocessos, és possible dividir un procés complex en diversos nivells, la qual cosa permet centrar-se en una àrea determinada d'un diagrama de procés.
- **Call Activity** o Activitat de trucada: És una activitat definida en un procés extern a la definició del procés actual. Permet crear una definició de procés que pot ser reutilitzada en altres definicions.

Gateways

Les *Gateways* o Portes d'entrada determinen quin camí s'empra a través d'un procés que controla el flux tant de fluxos de seqüències divergents com convergents. És a dir, una única porta podria tenir múltiples entrades i múltiples fluxos de sortida en els quals es permet o no l'entrada o la sortida. N'hi ha de diversos tipus, a continuació se n'explica tres dels més importants:

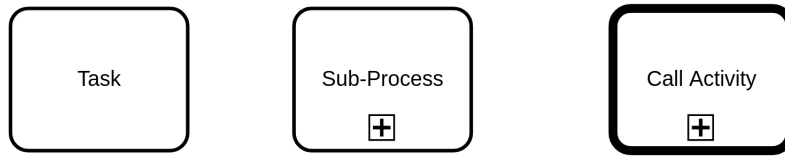


Figura 2.2: Representació dels diferents tipus d'Activitats. Font: Elaboració pròpia.

- **Exclusive Gateway** o Porta exclusiva: Se segueix un únic camí.
- **Parallel Gateway** o Porta paral·lela: Se segueixen tots els camins.
- **Inclusive Gateway** o Porta inclusiva: Se segueix un camí o més.



Figura 2.3: Representació dels diferents tipus de *Gateways*. Font: Elaboració pròpia.

Connectors

Els elements anteriors es connecten entre si en un diagrama per crear l'estructura bàsica d'un procés empresarial. Hi ha tres objectes de connexió que proporcionen aquesta funció. Aquests connectors són:

- **Sequence Flow** o Flux de seqüències: S'utilitza per mostrar l'ordre (la seqüència) en què es realitzaran les activitats en un procés.
- **Message Flow** o Flux de missatges: Representa el flux d'informació a través dels límits de l'organització. Els grups, les activitats i els esdeveniments de missatges es poden associar amb el flux de missatges. El flux de missatges es pot personalitzar amb un sobre que mostri el contingut del missatge.
- **Association** o Associació: Les anotacions permeten afegir més informació al diagrama que ajudi a documentar el procés.

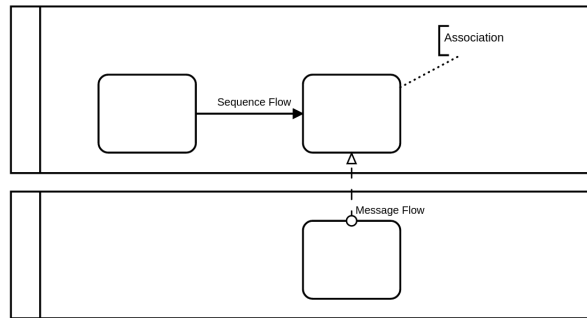


Figura 2.4: Representació dels diferents tipus de Connectors. Font: Elaboració pròpia.

3

Justificació

Donat que el modelatge de processos amb l'estàndard BPMN és molt específic l'estat de l'art no és molt extens. Tot i això, es poden trobar diverses aplicacions similars, encara que no fan ben bé el que es pretén en aquest treball. Aquests projectes fets anteriorment els podem classificar entre els que transformen d'un diagrama BPMN a text natural i els que ho fan a l'inrevés.

Abans, desenvolupar una eina que automatitzés qualsevol procés requeria molt esforç i dedicació, ja que, s'havien de tenir en compte totes les possibles combinacions en els possibles *inputs*. Això feia que en certs camps, com ara el de convertir un text natural, fos pràcticament impossible d'aconseguir. Per sort, mètodes com el *Deep Learning* permeten generar contingut de forma ràpida i eficient. Però, per aconseguir un model funcional es requereix un conjunt de dades molt gran i divers, i això no sempre és fàcil d'aconseguir.

Per a solucionar el problema plantejat en aquest treball no disposem d'un bon *dataset* que ens permeti generar un model de *Deep Learning* fiable. És per això, que la part principal d'aquest treball consisteix a crear una aplicació que generi un conjunt de dades variat i amb prou registres.

Això es farà convertint diversos diagrames a un llenguatge natural que sigui entenedor per qualsevol persona sense coneixement de BPMN. En aquest cas s'ha decidit que es farà servir l'anglès, ja que és un dels més utilitzats al món.

Per a fer aquesta primera part, hi ha un treball del 2016 titulat *Transformant Models BPMN a Llenguatge Natural* [7] que fa quelcom semblant. En el treball es llegeix el fitxer BPMN, després es crea un RSPT (Refined Program Structure Tree) per a guardar la informació necessària per solucionar el problema i també es fa servir una xarxa de Petri per a extreure l'estructura del procés. Finalment, es fa la generació del text analitzant cada element i concatenant els diferents missatges seguint l'estructura dels arbres creats anteriorment.

La segona part del projecte consisteix a entrenar una xarxa neuronal que sigui capaç d'interpretar una descripció en llenguatge natural i que ho converteixi en un model BPMN. Per a fer això, es farà servir el dataset que hem creat a la primera part. De treballs anteriors i aplicacions existents que converteixin de text a diagrames n'hi ha algunes que s'esmenten a continuació. Però cap d'elles fa servir un model de *Deep Learning*. Per tant, en aquest cas caldrà entrenar el model des de zero.

Process Talks [8] és una aplicació web que permet modelar processos a partir d'una explicació informal, sigui escrivint o parlant. L'eina fa servir regles i patrons que analitzen l'estructura sintàctica de les frases i treuen informació de qui realitza cada acció, o quina acció precedeix una altra. Aquesta implementació és costosa de mantenir i no és escalable, ja que, és complicat traspasar-la a nous idiomes. Per això, l'objectiu del projecte és crear una alternativa basada en *Machine Learning*.

A part d'aquesta aplicació web, també podem trobar un treball de recerca anomenat *Process Model Generation from Natural Language Text* [9]. Podem veure que es combinen algunes eines existents en el processament del llenguatge natural per a obtenir uns millors resultats. L'avaluació mostra que per a un conjunt de 47 parells de text-model de la indústria i els llibres de text, generen de mitjana el 77% dels models correctament.

Com s'ha pogut veure hi ha poques implementacions que serveixin de cara a plantejar una solució. Tanmateix, s'intentarà aprofitar el màxim de recursos disponibles per a reduir costos i temps.

4

Abast

4.1 Objectius

El principal objectiu del projecte és crear un conjunt de dades anotades que relacionin textos amb els models formals que els descriuen, per tal que puguin ser usades per entrenar sistemes neuronals de Deep Learning per fer aquesta tasca.

Com ja s'ha dit anteriorment a l'apartat de Justificació, el projecte es pot dividir en dos parts. La primera és la **Generació del Dataset** i la segona és la d'**Entrenar el model**. En la figura 4.1 es pot veure les diferents parts del projecte i els subobjectius que té cadascuna d'elles.

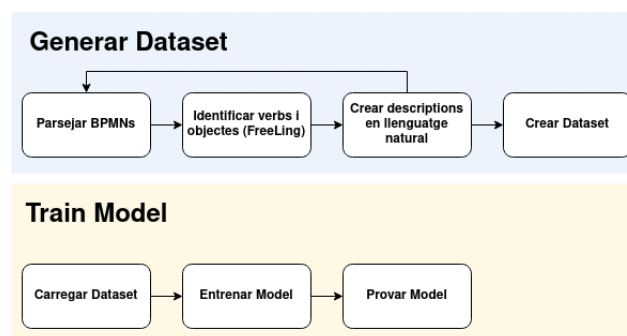


Figura 4.1: Subobjectius de cada part del projecte. Font: Elaboració pròpia.

En fase de generació del dataset el que es busca és tenir un conjunt de dades prou gran per a poder entrenar després el model.

Per a generar les dades es crearà un programa que llegirà (en anglès *parse*), després s'identificaran els verbs i la resta de paraules utilitzant un software d'anàlisi lingüístic anomenat *Freeling* [10]. Un cop s'hagi fet l'anàlisi sintàctica de les tasques es procedirà a crear les descripcions en llenguatge natural, utilitzant un software de generació de llenguatge natural anomenat *SimpleNLG*[11].

Un cop s'hagi desenvolupat el programa anterior, simplement s'haurà de repetir el procés de generació de text per a cada un dels models BPMN i emmagatzemar el conjunt de les dades en un format adequat per a poder-ho utilitzar després.

La segona part del projecte és la d'entrenar un model de *Deep Learning*. En aquest pas haurem de carregar el dataset generat anteriorment, entrenar el model i provar que el model ens retorna el resultat que volem. Si no hem obtingut un resultat satisfactori, s'ha de polir el model canviant paràmetres fins que estiguem contents amb els resultats.

4.2 Possibles obstacles

És molt comú que quan es comença a fer un projecte es trobin obstacles i problemes que no s'havien plantejat. Quan això passa pot ser que el projecte s'endarrereixi molt. Per això, s'ha d'intentar tenir en compte el conjunt de riscos que poden sorgir.

Un dels obstacles que ens podem trobar és que la generació de textos automàtica no doni uns resultats prou convincents. La informació disponible per a realitzar un projecte és un factor molt important a tenir en compte, i el fet que no puguem trobar prou informació sobre un tema és un risc que també hem de tenir en compte. També s'ha de tenir present que el temps és limitat, per tant, és possible que no sempre es puguin assolir tots els objectius marcats. A més, hem de tenir presents els següents obstacles que apareixen quan es fa un model de *Deep Learning*.

- El dataset no és prou divers.
- No es tenen gaires models BPMN per a crear el dataset.
- Obtenir bons resultats amb un model de ML és complicat, ja que hi intervenen moltes variables.
- Un bon model pot requerir un ús de recursos molt gran.
- Un bon model pot requerir un temps d'entrenament massa llarg.

5

Metodologia

5.1 Mètode Àgil

La metodologia escollida per al desenvolupament del software d'aquest projecte és l'*Agile* [12]. La base d'aquest conjunt de principis i pràctiques és permetre una planificació iterativa, i això implica descobrir requisits i desenvolupar solucions mentre s'avança en el desenvolupament. A cada iteració del cicle de vida es planifica, es dissenya, es desenvolupa, es testeja, es revisa i si convé es torna a iniciar la iteració. En el cas que se superi el control, llavors l'aplicació es publica.

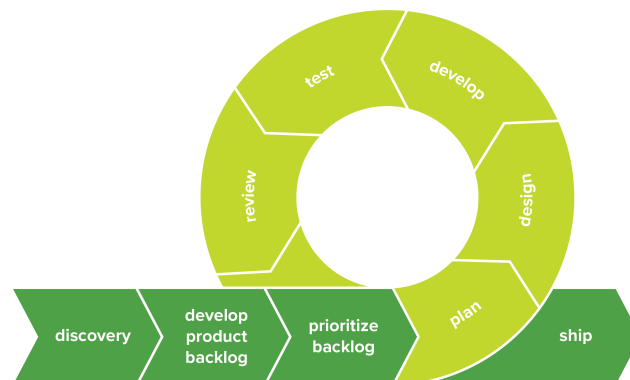


Figura 5.1: Metodologia Agile. Font: csgsolutions.

5.2 Eines

En aquest apartat s'explica el conjunt d'eines que ens permetran aplicar la metodologia *Agile* al nostre projecte.

Per començar necessitem un controlador de versions [13], això és un software que ens permet guardar i gestionar una còpia de les diferents versions del nostre programa. S'ha escollit *Git* [14] en combinació amb *GitHub* [15], què és un servei que web què permet fer ús del control de versions en línia.

Per a comprovar que el desenvolupament de les tasques es fa correctament fem servir un servei web d'organització i gestió de tasques anomenat *Todoist* [16]. Aquest software permet definir les diferents activitats que cal accomplir en cada part del projecte. També es pot definir una data límit per a cadascuna d'elles. Així doncs, ens permet saber si portem el desenvolupament al dia.

5.3 Validació

El software es revisa cada setmana amb els directors del treball, de manera que en cas d'haver-hi algun error es pot decidir com es procedeix de forma ràpida. A més, aquesta metodologia també permet proposar canvis en el disseny del programa, d'aquesta manera es pot afegir o eliminar funcionalitats en funció del temps restant per a l'entrega.

6

Planificació temporal

Un dels aspectes més importants a l'hora de dur a terme un projecte és el temps del qual es disposa. En el moment en què es redacta aquest document la data per a la finalització del treball encara no està definida, de totes maneres, fixem el límit per acabar el projecte a principis del mes de gener del 2022. El treball es va començar el 10 de juliol del 2021. Això vol dir que es disposen de 184 dies en total per a fer el treball, inclosos caps de setmana i festius. S'estima que en un Treball de Fi de Grau (TFG) la dedicació total és de 540 h. Això implica que s'haurien de destinar unes 3 h al dia per a complir amb els objectius establerts. Evidentment no cada dia es treballarà aquestes h, però sí que s'haurien de fer unes 21 h setmanals.

6.1 Definició de tasques

En aquest projecte s'ha decidit dividir la feina a fer en un conjunt de tasques. Aquesta divisió permet tenir un control més gran del projecte. De manera que es pugui saber l'inici, el final i les dependències que té cada tasca. Això ens permet saber en tot moment si el projecte està endarrerit o no. I en cas que ho estigui, prendre les mesures oportunes.

A continuació s'expliquen els recursos necessaris per al projecte i totes les tasques i subtasques que s'han definit per al treball.

6.1.1 Recursos necessaris

Hem determinat que els recursos necessaris per a desenvolupar el projecte són:

Humans - R1

Els recursos humans necessaris són tots dos codirectors del projecte Josep Carmona i Lluís Padró, la tutora del GEP Olga Pons i l'autor del treball o programador Guillem Pla.

Maquinari - R2

El maquinari que fem servir és un ordinador portàtil de gamma mitjana. Aquest ens permet tant documentar com programar com entrenar el model. A més, ens deixa moure allà on vulguem i per tant facilitar les reunions presencials.

El portàtil utilitzat és un *Acer Aspire 3*.

Programari - R3

El software que es preveu utilitzar el *GitHub*(controlador de versions), *Todist*(gestió del projecte), *PyCharm*(programació de projectes en Python) i *Google Collab*(executar codi al núvol) i *Tensorflow*(llibreria Python de machine learning).

6.1.2 Gestió del Projecte - GP

La primera part del projecte és la de gestionar de forma correcta el projecte. En aquest apartat s'hi engloben les gestions no tècniques que cal fer per a engegar el projecte. El temps total per a realitzar aquesta part s'estima en 90 hores. Aquestes són la contextualització i abast, la planificació temporal, el desenvolupament d'un pressupost, la realització de l'informe de sostenibilitat, l'elecció de la metodologia i la instal·lació del programari necessari. A continuació s'expliquen en més detall.

Contextualització i abast - GP-C

En aquesta tasca es defineixen els fonaments del projecte. Dins d'aquesta s'hi inclou la tria del tema, l'objectiu principal a resoldre, l'explicació de per què és necessari fer aquest treball i l'abast d'aquest. Per tant, una part de la feina a realitzar consisteix en fer reunions amb els directors del treball i una altra en documentar les decisions preses. S'estima que es realitza en unes 15 hores.

Planificació temporal - GP-T

Aquí es tracta de dividir el treball en tasques i d'establir un calendari per a cadascuna d'elles. Això es fa per a poder tenir, en tot moment, el control de la feina feta i la que falta per fer. S'estima que són necessàries unes 13 hores.

Metodologia - GP-M

En aquesta tasca el que es fa és estudiar el conjunt de metodologies de treball disponibles i fer una tria adequada d'una d'elles. Això permetrà seguir el calendari d'una forma més acurada. S'estima que la tasca de Metodologia es pot dur a terme en 13 hores.

Pressupost - GP-P

En aquest apartat es crea un informe on s'expliquen els termes econòmics del treball. D'acord amb les hores realitzades i amb els recursos emprats es fa un càlcul aproximat sobre el cost que té crear el projecte. Aquesta tasca es pot realitzar en unes 13 hores.

Informe de sostenibilitat - GP-S

Aquesta tasca és molt semblant a l'anterior, però en aquest cas, el que s'estudia és l'impacte mediambiental i social que el treball pot tenir. Aquesta tasca s'hauria de desenvolupar en 13 hores.

Instal·lació de programari - GP-I

Abans de començar a programar és vital decidir quines eines de software es faran servir. Per això, en aquesta tasca el que es fa és un estudi d'un conjunt d'eines disponibles al mercat i posteriorment es decideix quines es volen fer servir. Per a decidir entre un programa o un altre es té en compte la facilitat d'ús, si ja s'ha fet servir prèviament, la corba d'aprenentatge, la gratuïtat del programa i si és de codi obert. Un cop ja s'ha fet la tria només queda instal·lar totes les eines escollides. S'estima que es pot tardar unes 20 hores.

6.1.3 Generació del dataset - GD

Un cop tenim definit el projecte a realitzar (GP-C 6.1.2) ja podem començar amb la part de desenvolupament. La primera part del desenvolupament consisteix a generar un conjunt de dades que ens permeti després entrenar un model de Deep Learning. Aquest conjunt de dades constarà d'un conjunt de parelles formades per un model BPMN i un text descriptiu del model. En aquest apartat s'espera invertir unes 130 hores en total. Per a aconseguir això necessitem dur a terme les tasques següents.

Obtenir models BPMN - GD-O

Primer de tot, necessitem obtenir un conjunt de models BPMN que ens serveixi de base per a poder desenvolupar i testear el nostre programa. D'aquests diagrames BPMN se'n poden trobar gratuïtament a internet, per tant, el que cal fer és una cerca a la xarxa i descarregar-los. També és convenient comprovar que els arxius estan modelats correctament i es poden fer servir. S'estima que la durada d'aquesta tasca és d'unes 10 hores. Aquesta tasca no es pot iniciar fins que no s'hagi contextualitzat el projecte (GP-C).

Parsejar BPMN - GP-P

Un cop tenim una sèrie de models descarregats, procedim a llegir l'arxiu que els conté, és a dir, a parsejar-los. Per a llegir-los caldrà recórrer el fitxer *.bpmn* i emmagatzemar cada element útil en un arbre. S'estima que la durada d'aquesta tasca és d'unes 14 hores.

Aquesta tasca no es pot iniciar fins que no s'hagi finalitzat la tasca d'obtenir models BPMN (GP-O).

Identificar verbs - GP-I

La tasca anterior ha permès obtenir el conjunt d'elements que conformen el model. Aquests elements sovint tenen una curta frase explicant en què consisteix la tasca del model. Aquesta frase s'utilitzarà més endavant per a generar un text natural, però abans es necessita l'anàlisi sintàctica d'aquesta frase. Per a fer-ho, farem servir la llibreria *Freeling* que ens retornarà el tipus de cada paraula de la frase. Aquesta informació l'emmagatzemarem per a fer-la servir més endavant. S'estima que la durada d'aquesta tasca és d'unes 31 hores. Aquesta tasca no es pot iniciar fins que no s'hagi finalitzat la tasca de parsejar BPMN (GP-P).

Crear descripcions - GP-D

En aquesta tasca es crea una descripció en llenguatge natural a partir d'un diagrama BPMN. Per a fer-ho cal utilitzar la informació emmagatzemada anteriorment i fer-la servir d'input. Per a crear aquest programa es fa servir una llibreria de generació de llenguatge natural anomenada SimpleNLG. Aquest programa ens retornarà la descripció del model introduït. S'estima que la durada d'aquesta tasca és d'unes 47 hores. Aquesta tasca no es pot iniciar fins que no s'hagi finalitzat la tasca d'identificació de verbs (GP-I).

Crear dataset - GP-C

Al final, utilitzarem el programa creat en la tasca anterior per a processar múltiples diagrames. D'aquesta forma obtindrem per a cada model una descripció que ens servirà per a crear el dataset. Un cop tinguem estiguem satisfets amb el nombre de descripcions creades es procedeix a ajuntar-ho tot en un mateix arxiu. S'estima que la durada d'aquesta tasca és d'unes 28 hores. Aquesta tasca no es pot iniciar fins que no s'hagi finalitzat la tasca de creació de descripcions (GP-D).

6.1.4 Crear model Deep Learning - MD

La segona part del desenvolupament consisteix a crear un model de Deep Learning que utilitzi el dataset que hem creat anteriorment per a aprendre a descriure un model BPMN automàticament. En aquesta fase és les tasques d'entrenament i provar el model se superposen. Això passa perquè un cop es tingui un model entrenat s'ha de comprovar si funciona bé. En Machine Learning ja se sap que no s'obté un model funcional de seguida, per tant, s'ha de validar el model i en cas que no s'obtingui el resultat desitjat s'haurà de generar un nou model. I així successivament. Per a realitzar el conjunt de tasques d'aquesta secció s'espera destinar unes 172 hores.

Carregar dataset - MD-C

Per a crear un model de Deep Learning necessitem carregar primer el dataset desitjat. Així doncs, s'haurà de llegir el fitxer que conté les dades amb una llibreria adequada. En principi sembla una tasca fàcil, però la lectura de fitxers sempre pot donar problemes. I més quan el conjunt de dades ha sigut generat per nosaltres mateixos. S'estima que la durada d'aquesta tasca és d'unes 36 hores. Aquesta tasca no es pot iniciar fins que no s'hagi finalitzat tot l'apartat de Gestió de Projectes (GP) i l'apartat de Generació del Dataset (GD).

Entrenar model - MD-E

A continuació ja podem entrenar el model usant les dades carregades. D'entrada, en aquesta tasca s'ha de definir quina arquitectura es vol fer servir en el model. Després es definiran uns paràmetres i es procedirà a entrenar el model. Un cop ja s'hagi provat i no s'hagi obtingut uns bons resultats, en aquesta tasca es procedeix a canviar els paràmetres amb la intenció de millorar la generació final del text. Aquesta tasca per tant es pot allargar molt. S'estima que la durada d'aquesta tasca és d'unes 72 hores. Aquesta tasca no es pot iniciar fins que no s'hagi finalitzat la tasca de carregar el dataset (MD-C).

Provar model - MD-P

Aquesta tasca consisteix a comprovar si els resultats que dona el model de Deep Learning són prou bons. Com que no es tracta d'un procés de classificació automàtic validar que la sortida obtinguda és un procés manual i per tant, costós. A més, és bastant subjectiu. Per evitar-ho es crearan uns tests que validin automàticament certs aspectes del text generat. Aquesta tasca es durà a terme pràcticament alhora que la d'entrenar el model. S'estima que la durada d'aquesta tasca és d'unes 72 hores. Aquesta tasca no es pot iniciar fins que no s'hagi finalitzat la tasca d'entrenar un model (MD-E).

6.1.5 Documentació - DO

La documentació és una tasca molt important del treball, ja que és la que permet explicar com s'ha organitzat, realitzat i quins resultats s'ha obtingut. Per això, s'ha de dur a terme paral·lelament durant tot el desenvolupament. En aquest apartat s'espera invertir unes 120 hores en total.

Memòria - DO-M

Aquesta tasca es farà alhora que la resta del projecte i contextualitzarà el treball, es detallaran tots els passos que s'han seguit i es justificarà cada decisió presa. S'estima que la durada d'aquesta tasca és d'unes 90 hores.

Presentació - DO-P

En aquesta tasca es crearà una presentació que contindrà un breu resum del projecte elaborat. Aquesta serà un ajut visual de cara a la presentació que es faci d'aquest treball d'avant del tribunal que avaluï el projecte. Dins d'aquesta tasca també s'hi ha de tenir en compte el temps per a preparar l'explicació oral. S'estima que la durada d'aquesta tasca és d'unes 30 hores. Aquesta tasca no es pot iniciar fins que no s'hagi finalitzat el treball (GP, GD, MD, DO-M).

Codi	Nom Tasca	Temps Estimat	Dependències Temporals	Recursos
GP-C	Contextualització i abast	15 h	-	R1, R2
GP-T	Planificació temporal	13 h	GP-C	R1, R2
GP-M	Metodologia	13 h	GP-C	R1, R2
GP-P	Pressupost	13 h	GP-C	R1, R2
GP-S	Informe de sostenibilitat	13 h	GP-C	R1, R2
GP-I	Instal·lació programari	20 h	-	R1, R2
GD-O	Obtenir models BPMN	10 h	GP-C	R1, R2, R3
GD-P	Parsejar BPMN	14 h	GD-O	R1, R2, R3
GD-I	Identificar verbs	31 h	GD-P	R1, R2, R3
GD-D	Crear descripcions	47 h	GD-I	R1, R2, R3
GD-C	Crear dataset	28 h	GD-C	R1, R2, R3
MD-C	Carregar dataset	36 h	GP, GD	R1, R2, R3
MD-E	Entrenar model	72 h	MD-C	R1, R2, R3
MD-P	Provar model	64 h	MD-E	R1, R2, R3
DO-M	Memòria	90 h	-	R1, R2
DO-P	Presentació	30 h	GP, GD, MD, DO-M	R1, R2

Taula 6.1: Taula de les tasques. Font: Elaboració pròpia.

6.2 Estimacions i Gantt

En aquest apartat utilitzarem un diagrama de Gantt per a visualitzar l'ordre i duració de cadascuna de les tasques descrites anteriorment. Un diagrama de Gantt [17] és un tipus de gràfic que il·lustra el calendari d'un projecte. Aquest diagrama mostra les tasques que s'han d'executar en l'eix vertical i en l'eix horitzontal mostra el moment d'inici, el moment de finalització i la duració de cada tasca. En el nostre cas hem decidit dividir els 184 dies en 27 setmanes. Com hem comentat anteriorment en l'apartat de Planificació temporal 6 en cada setmana es disposa de 21 h per a executar les diferents tasques. A continuació es mostren dos diagrames de Gantt, un per a la primera part del projecte, les primeres 13 setmanes, i un per a la segona part del projecte, corresponent a les següents 14 setmanes.

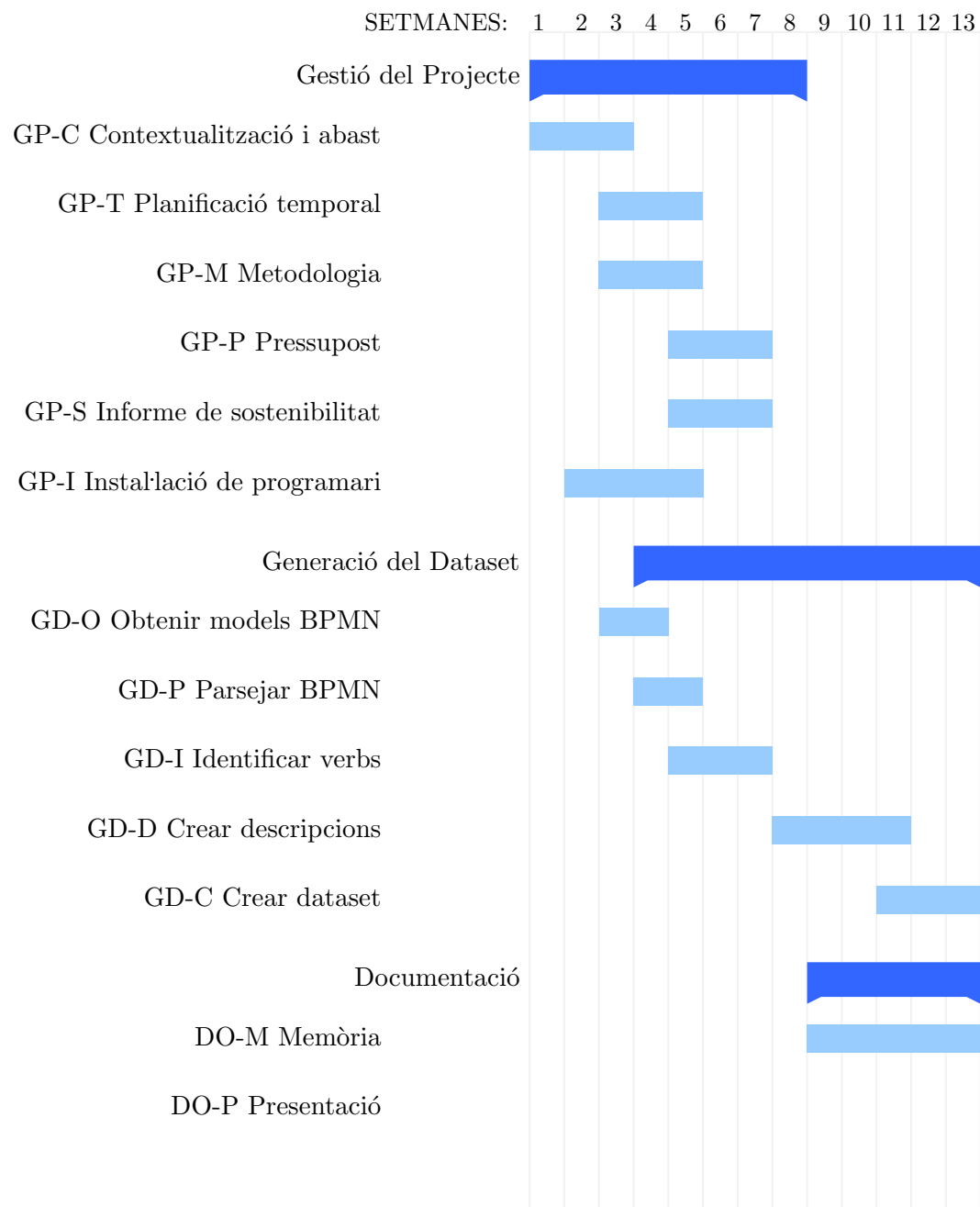


Figura 6.1: Primera part del diagrama de Gantt del projecte. Font: Elaboració pròpia.

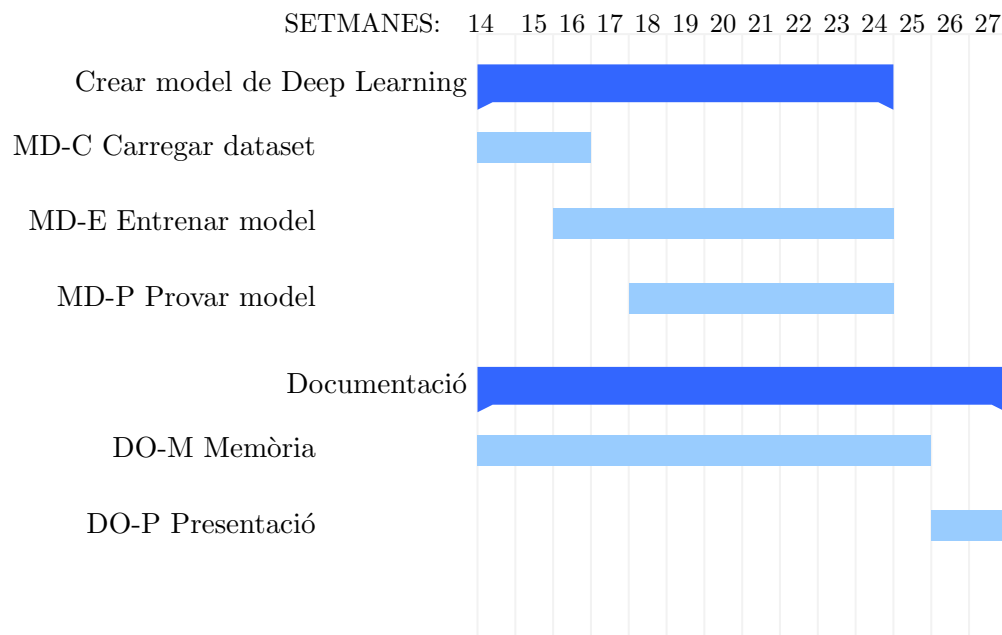


Figura 6.2: Segona part del diagrama de Gantt del projecte. Font: Elaboració pròpia.

6.3 Plans alternatius i obstacles

En tot projecte s'ha de tenir previst que puguin aparèixer inconvenients. En conseqüència cal pensar bé quins problemes poden sorgir en cada fase del treball i proposar una solució alternativa. A continuació s'explica quins plans es poden seguir en cas que apareguin alguns obstacles.

- Un dels principals problemes que sorgeix a quasi tots els projectes és la manca de temps. Per abordar aquest obstacle s'ha decidit que no cal crear cap tasca alternativa, sinó que s'intentarà treballar més hores de les proposades per poder arribar a temps a la data límit.

També s'ha tingut present això a l'hora d'assignar la càrrega a cada tasca. Podem veure que el temps estimat total per a la realització del projecte és d'unes 509 hores, però en TFG s'haurien de dedicar unes 540 hores. Llavors tenim unes 31 hores que es poden realitzar de més i seguir acabant a temps el pla.

A més, en l'anàlisi de temps dedicat a cada tasca també s'ha tingut en compte que entrenar i provar el model no sortirà bé de seguida. Per tant, els temps estimats ja tenen en compte que es pugui tardar a obtenir un resultat desitjable.

- Un altre dels problemes que pot sorgir és que la informació sobre el tema del treball no sigui gaire extensa, això suposa un problema per què costa més entendre

com s'ha de procedir. La solució és espremer al màxim la informació disponible i aprendre a partir de la investigació pròpia.

- Un altre problema que pot sorgir és que el dataset creat no sigui prou divers o que no es tinguin gaires models BPMN per a crear el dataset. Si un d'aquests casos apareix, implicaria que mai es podria obtenir un bon model de Deep Learning. Com que això no és tolerable s'hauria de crear una nova tasca anomenada Crear models BPMN inventats (GD-O). Es tractaria d'usar *data augmentation* per a obtenir més models BPMN i així intentar aconseguir un dataset més divers. Evidentment, això és complicat de realitzar i no garanteix uns bons resultats, però s'hauria de provar. La realització d'aquesta tasca podria durar unes 15-25 hores.
- L'últim possible obstacle que s'ha analitzat és que no es disposi de recursos suficients per a entrenar el model. En aquest cas, la solució proposada és executar el codi en servidors al núvol amb més velocitat de processament que les nostres màquines en local. Aquests podran ser tant de pagament com gratuïts. S'intentarà utilitzar sempre l'opció més econòmica. Si cal fer servir aquests serveis no comportarà un temps afegit, ja que la posada a punt d'aquests servidors és molt ràpida i intuïtiva.

Pressupost

En aquest apartat s'elabora una anàlisi econòmica del projecte amb la finalitat de decidir si és factible dur-lo a terme. Hem d'analitzar i estimar les següents classes de costos: costos de personal, costos de les tasques, maquinari, programari i costos d'espai de treball. També s'ha de tenir en compte els possibles contratemps que puguin sorgir i fer una estimació del que poden costar. A continuació es detallen les classes esmentades.

7.1 Costos de personal

Abans de res, hem de definir quins rols hi ha involucrats en el projecte. En aquest cas, hi ha el càrrec de Cap de Projecte, que és la persona encarregada de planificar, supervisar i dirigir-lo durant la realització d'aquest. És un rol compartit entre els codirectors del TFG i l'autor d'aquest. També existeix el rol de programador, que és la persona que dissenya, escriu i depura el codi del projecte. En aquest cas el rol només l'executa l'autor del TFG.

A la taula 7.1 veiem el sou brut, el que s'ha de pagar de Seguretat Social (SS) i la retribució total de cadascun dels càrrecs esmentats anteriorment. Els sous bruts han sigut extrets de la pàgina [18].

Rol	Sou brut	SS	Retribució
Cap de projecte	22,5 €/hora	6,75 €/hora	29,25 €/hora
Programador	12 €/hora	3,6 €/hora	15,6 €/hora

Taula 7.1: Rols del projecte i la facturació per hora. Font: Elaboració pròpia.

7.2 Costos per tasca

El temps de desenvolupar cada tasca és diferent, per tant cal calcular per a cada una de les activitats el seu cost. El temps de cada tasca està descrit a l'apartat de Planificació temporal 6. També hem de considerar que cada tasca pot ser desenvolupada pel Cap

del Projecte o pel Programador. Les activitats associades a la Gestió del Projecte (GP) 6.1.2 i a la Documentació 6.1.5 les durà a terme íntegrament el Cap del Projecte. Les tasques de Generació del Dataset 6.1.3 les desenvoluparà el Programador. En canvi, les tasques d'Entrenament del Model 6.1.4 les faran el Cap del Projecte i el Programador, això sí, el Programador és el que hi ha d'invertir més temps. En aquestes tasques el Cap del Projecte simplement ajudarà a dissenyar el model i decidirà si és correcte o no.

Codi	Nom Tasca	Cap del projecte	Programador	Cost
GP-C	Contextualització i abast	15 h	0 h	438,75 €
GP-T	Planificació temporal	13 h	0 h	380,25 €
GP-M	Metodologia	13 h	0 h	380,25 €
GP-P	Pressupost	13 h	0 h	380,25 €
GP-S	Informe de sostenibilitat	13 h	0 h	380,25 €
GP-I	Instal·lació programari	0 h	20 h	312,00 €
GD-O	Obtenir models BPMN	0 h	10 h	156,00 €
GD-P	Parsejar BPMN	0 h	14 h	218,40 €
GD-I	Identificar verbs	0 h	31 h	483,60 €
GD-D	Crear descripcions	0 h	47 h	733,20 €
GD-C	Crear dataset	0 h	28 h	436,80 €
MD-C	Carregar dataset	0 h	36 h	561,60 €
MD-E	Entrenar model	16 h	56 h	1.341,60 €
MD-P	Provar model	12 h	52 h	1.162,20 €
DO-M	Memòria	90 h	0 h	2632,50 €
DO-P	Presentació	30 h	0 h	877,50 €

Taula 7.2: Taula del cost per cada tasca. Font: Elaboració pròpia.

En la taula 7.2 hi trobem el codi i nom de totes les tasques que s'han de realitzar. També hi trobem el temps que hi dediquen tant el Cap del Projecte com el Programador. I finalment també hi trobem el cost total de cada tasca, aquest s'ha calculat en funció del preu per hora calculat a la taula 7.1. En total el cost de totes les tasques és de 10.875,15 €.

7.3 Costos genèrics

En aquest apartat explicarem els costos genèrics que tenim al projecte. Aquests costos són la suma dels costos de maquinari, de programari i del teletreball. A continuació es detallen cadascun d'ells.

7.3.1 Maquinari

En tot projecte informàtic es necessita un maquinari per a dur-lo a terme. En aquest cas s'ha fet servir un monitor *Newskill Icarus RGB IC27QRS 27"*, un portàtil *Acer Aspire 3 A315-51-59SU*, un ratolí *Logitech G603* i un teclat *Logitech G213*. A la taula 7.3 veiem

el seu preu de mercat l'octubre del 2021 i també el seu cost amortitzat. Per a calcular el cost amortitzat s'ha de tenir en compte que el maquinari té un temps de vida estimat de 4 anys. Per a fer el càlcul de les hores reals que permet treballar cada eina hem de tenir en compte que un any disposa d'uns 231 dies hàbils (251 dies laborables menys 20 dies de vacances) i que 1 dia hàbil compta de 8 hores. Com ja hem explicat a l'apartat de Planificació 6, calculem que farem servir aquestes eines unes 540 hores. Per tant, la fórmula que farem servir és la següent.

$$4 \text{ anys} \cdot 231 \text{ dies laborables} \cdot 8 \text{ hores} = 7392 \text{ hores hàbils}$$

$$\frac{540 \text{ hores de projecte}}{7392 \text{ hores hàbils}} = 0,073$$

Per tant, multipliquem el preu de mercat de cada producte per 0,073 per a obtenir el cost amortitzat.

Nom del maquinari	Preu de mercat	Cost amortitzat
Monitor	239 €	17,45 €
Portàtil	519 €	37,89 €
Ratolí	49 €	3,50 €
Teclat	62 €	4,53 €
Total	869 €	63,37 €

Taula 7.3: Costos del maquinari. Font: Elaboració pròpia.

7.3.2 Programari

Actualment al mercat hi ha disponible una gran quantitat de programari. Aquest pot ser gratuït o de pagament. Normalment es pot trobar software gratuït de qualitat que ens permet realitzar la tasca i a més estalviar diners. Per a dur a terme aquest treball s'ha intentat reduir al màxim els costos de programari. És per això que el llistat de programes que mostrem a la taula 7.4 és totalment gratuït. Alguns dels serveis que utilitzarem també tenen una versió de pagament, però en principi amb la versió més bàsica d'aquests ens és suficient. També és important comentar que el *PyCharm Edu* és només gratuït per a estudiants, com que l'autor d'aquest treball ho és el podem fer servir, si no ens costaria 19,90 € al mes.

7.3.3 Teletreball

Com que estem en pandèmia i és millor evitar els espais públics aquest treball es fa íntegrament a casa. Per tant, també s'ha de calcular quant costa el fet d'estar teletreballant a casa. Això implica el mobiliari (taules, cadires, etc.), la llum, el lloguer i altres despeses. Alguns sindicats laborals [19] ja estan calculant quant costa el fet de treballar des de casa, i alguns d'ells han arribat a la conclusió que aquestes despeses arriben a uns 21 € al mes. Per tant, si el projecte es desenvolupa en 7 mesos, el cost total és de 147 €.

Nom programari	Cost mensual	Mesos utilitzats	Cost total
Python	0 €	6	0 €
PyCharm Edu	0 €	6	0 €
Overleaf	0 €	7	0 €
Clockify	0 €	7	0 €
Github	0 €	7	0 €
Tensorflow	0 €	4	0 €
Google Colab	0 €	4	0 €

Taula 7.4: Costos pel programari. Font: Elaboració pròpia.

7.4 Total dels costos per personal i genèrics

En la taula 7.5 podem veure el total dels costos del personal i els costos genèrics calculats anteriorment.

Concepte	Cost
Personal	10.875,15 €
Maquinari	63,37 €
Programari	0,00 €
Teletreball	147,00 €
Total	11.085,52 €

Taula 7.5: Costos de personal i genèrics. Font: Elaboració pròpia.

7.4.1 Contingència

Cal tenir present que durant el projecte surtin dificultats i adversitats. Aquestes poden fer que el cost final del projecte s'encareixi. Per tant, hem d'estar preparats per si això passa. Preparem una partida de contingència, o marge de seguretat, que ens permetrà cobrir els imprevistos no anticipats. Es calcula com un percentatge del total del pressupost.

En els projectes de desenvolupament software s'acostuma a reservar entre un 10% i un 20% del pressupost. Com que tenim un nivell de detall del pressupost òptim, només destinarem el 10%. Per tant el cost total de la contingència el calculem així:

$$\text{Cost total} \cdot 10\% = 11.085,52 \cdot 10\% = 1.108,55 \text{ €}$$

7.4.2 Imprevistos

En darrer terme també s'ha de considerar les activitats que formen part dels plans alternatius. Aquestes activitats no es costegen totalment, sinó que es fa segons el percentatge estimat de risc que tinguin cadascuna d'elles. A la taula 7.6 veiem els possibles

imprevistos que poden sorgir, l'increment de cost que suposarien i el cost que reservem finalment.

Imprevist	Cost	Risc	Total
Nou ordinador	500,00 €	5%	25,00 €
Increment del temps d'entrenament (20h)	339,30 €	20%	67,86 €
Google Collab Pro (5 mesos)	49,95 €	10%	4,99 €
Total	889,25 €		97,86 €

Taula 7.6: Taula amb els costos dels imprevistos. Font: Elaboració pròpia.

7.5 Cost total del projecte

En la taula 7.7 calcularem el cost total d'aquest.

Concepte	Cost
Tasques	10.875,15 €
Maquinari	63,37 €
Programari	0,00 €
Teletreball	147,00 €
Contingència	1.108,55 €
Imprevistos	97,86 €
Total	12.291,93 €

Taula 7.7: Taula amb els costos totals del projecte. Font: Elaboració pròpia.

7.6 Control de gestió

Per a evitar sorpreses en el cost s'ha de realitzar un control dels costos. Es tracta d'avaluar les possibles desviacions que tinguem durant el transcurs del projecte. Compararem els costos reals que hàgim tingut al treball amb el pressupostat. En cas de detectar una desviació hem de saber en quines activitats o etapes ha ocorregut, hem d'explicar el perquè ha passat i de quant ens hem desviat.

En cas que es produeixin desviacions en tasques mencionades a l'apartat d'imprevistos 7.4.2 s'utilitzarà la partida destinada a cobrir aquests contratemps. Si no fos suficient o les desviacions fossin d'una tasca que no hem tingut en compte hauríem de fer servir el fons de contingència. Les fórmules que utilitzarem per a fer el seguiment són les següents:

$$\frac{\text{cost estimat} - \text{cost real}}{\text{hores estimades} - \text{hores reals}}$$

Informe de sostenibilitat

Un cop respost el qüestionari del projecte EDINSOST som més conscients del grau de coneixement que tenim respecte a la sostenibilitat. Actualment es posa molt èmfasi en la sostenibilitat, el canvi climàtic i el medi ambient, això fa que tinguem coneixements teòrics generals sobre aquests temes. Malgrat tot, els coneixements pràctics necessaris per a acomplir un projecte de forma sostenible no els tenim.

Això implica que a l'hora de detectar que un element del treball no és sostenible ho podem fer amb certa facilitat. Però, no tenim prou coneixements per a mesurar-ho amb els indicadors adequats. També es fa difícil proposar alternatives més sostenibles, ja que ignorem les tècniques que es fan servir actualment.

Pel que fa a l'impacte social que un servei TIC pot tenir ens considerem capacitats per a avaluar el paper que juga aquest projecte en la societat. Tant en l'àmbit de l'accessibilitat, i la qualitat ergonòmica com per la justícia social, l'equitat i la diversitat. En aquest cas també ens considerem més capaços de proposar mesures eficaces per a millorar en aquests àmbits. Encara que també costa fer servir indicadors per a estimar com els projectes contribueixen a millorar el bé comú de la societat.

Respecte a la dimensió econòmica no comprenem gaire bé les parts econòmiques d'un projecte, ja que és un àmbit molt extens, però sí que tenim alguna idea.

I per finalitzar, considerem que som capaços de valorar els beneficis del treball en equip en les TIC, a més, coneixem les eines necessàries per a desenvolupar projectes amb més gent.

8.1 Dimensió econòmica

En l'apartat de Pressupost 7 s'ha detallat el cost del projecte i dels seus apartats per a tenir clar quines són les seccions amb una partida pressupostària més alta. Com hem explicat, s'ha intentat fer que tinguem un cost mínim. Això ha implicat fer servir programari gratuït i utilitzar maquinari que no sigui de gamma màxima.

El punt on més diners es destinen és en els costos personals. Evidentment, aquests no es poden reduir més, ja que el sou dels treballadors ha de ser just i les hores que han

de fer ja estan marcades per les característiques del TFG. En conseqüència creiem que el cost estimat per a la realització del projecte és just.

Pel que fa a les solucions actuals aquest projecte també pretén millorar-les en l'aspecte econòmic. Aquesta solució permetrà estalviar diners als usuaris d'aquest servei, ja que no es necessitarà tant personal qualificat per a realitzar els models BPMN.

8.2 Dimensió ambiental

La dimensió ambiental de realització del projecte s'estima que serà baixa. Això es deu al fet que no s'utilitza maquinari que consumeixi grans quantitats d'energia. A més, el treball es duu a terme plenament des de casa i per tant no es fan desplaçaments innecessaris que podrien tenir un cost ambiental elevat. Tampoc es produeixen residus, ja que el poc paper que es fa servir és o bé reciclat o bé reutilitzat.

Respecte a la seva vida útil aconseguirà reduir l'impacte que tenen les alternatives actuals. Amb l'ús de la solució plantejada en aquest projecte es redueix l'impacte ambiental, ja que es podrà reduir personal que treballi a fer models manualment. Això implica una reducció en desplaçaments, en espai de treball i en llum a les oficines.

8.3 Dimensió social

En l'àmbit personal creiem que el projecte ens aportarà coneixements tècnics en el sector del Deep Learning, coneixements sobre realitzar correctament un projecte d'enginyeria informàtica, capacitat per a entendre què són els models de processos empresarials, aprendrem a gestionar millor les tasques, millorarem la nostra capacitat per a executar projectes a temps i ens proporcionarà experiència a l'hora de programar.

Pel que fa a la utilitat d'aquest treball a nivell col·lectiu creiem que serà d'una gran utilitat a les empreses que facin servir BPMN, a més, ajudarà a reduir la quantitat de feina que els *Business Analysts* han de fer, i per tant, podran dedicar més temps a altres coses més necessàries. També pot ser que a la llarga es requereixin menys treballadors en una empresa fent aquesta feina, per a la persona que es dedica a això pot ser un problema, però al final la tecnologia sempre ha implicat canvis en totes les professions.

9

Solució plantejada

9.1 Eines utilitzades

9.2 Implementació

9.3 Solucions prèvies

9.4 Alternatives

9.5 Obstacles

10

Conclusions

11

Agraiments

Bibliografia

- [1] Institut d'Estudis Catalans. Diec2. <https://dlc.iec.cat/>. Accedit el 26/09/2021.
- [2] Ashley DiFranza. What does a business analyst do? <https://www.northeastern.edu/graduate/blog/what-does-a-business-analyst-do/>. Accedit el 26/09/2021.
- [3] Wikipedia. Business process model and notation. https://en.wikipedia.org/wiki/Business_Process_Model_and_Notation. Accedit el 20/08/2021.
- [4] Beyond BPMN. Beyond bpmn. <https://beyondbpmn.com/>. Accedit el 24/08/2021.
- [5] Marin Perez. The beginner's guide to using bpmn in business. <https://www.microsoft.com/en-us/microsoft-365/business-insights-ideas/resources/the-guide-to-using-bpmn-in-your-business>. Accedit el 24/08/2021.
- [6] Visual Paradigm. Bpmn activity types explained. <https://www.visual-paradigm.com/guide/bpmn/bpmn-activity-types-explained/>. Accedit el 16/10/2021.
- [7] Genís Martín Coca. Transformant models bpmn a llenguatge natural, 2016.
- [8] Josep Carmona, Lluís Padró, Marc Solé, and Sànchez-Ferrerres Josep. Process talks. <https://www.processtalks.com/>. Accedit el 26/09/2021.
- [9] Fabian Friedrich, Jan Mendling, and Frank Puhlmann. Process model generation from natural language text. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6741 LNCS:482–496, 2011.
- [10] Lluís Padró and Evgeny Stanilovsky. Freeling 3.0: Towards wider multilinguality. In *Proceedings of the Language Resources and Evaluation Conference (LREC 2012)*, Istanbul, Turkey, May 2012. ELRA.
- [11] Ehud Reiter. simplenlg. <https://github.com/simplenlg/simplenlg>.
- [12] Wikipedia. Agile software development. https://en.wikipedia.org/wiki/Agile_software_development. Accedit el 28/09/2021.

- [13] Wikipedia. Version control. https://en.wikipedia.org/wiki/Version_control. Accedit el 16/10/2021.
- [14] Wikipedia. Git. <https://en.wikipedia.org/wiki/Git>. Accedit el 16/10/2021.
- [15] Wikipedia. Github. <https://en.wikipedia.org/wiki/GitHub>. Accedit el 16/10/2021.
- [16] Todoist. Features. <https://todoist.com/features>. Accedit el 16/10/2021.
- [17] Wikipedia. Gantt chart. https://en.wikipedia.org/wiki/Gantt_chart. Accedit el 03/10/2021.
- [18] Glassdoor. Sueldo: Programador en barcelona. https://www.glassdoor.es/Sueldos/barcelona-programador-sueldo-SRCH_IL.0,9_IM1015_K010,21.htm?clickSource=searchBtn. Accedit el 10/10/2021.
- [19] CSIF, APT, UGT, CCOO, and Intersindical-CSC. Comunicado conjunto: Negociación de teletrabajo. <https://csifcapgemini.blogspot.com/2021/10/comunicado-conjunto-negociacion-de.html>. Accedit el 10/10/2021.