

# NLP per a la resolució de tiquets

Irina Moreno Lahoz

**Resum**– Resum del projecte, màxim 10 línies. ....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

**Paraules clau**– tiquets, NLP, Machine Learning, Intel·ligència Artificial, solució

**Abstract**– Versió en anglès del resum . ....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

**Keywords**– tickets, NLP, Machine Learning, Artificial Intelligence, solution



---

• E-mail de contacte: irina.moreno@autonoma.cat  
• Treball tutoritzat per: Débora Gil Resina (Ciències de la Computació)  
• Curs 2022/23

## 1 INTRODUCCIÓ

AQUEST treball s'ha desenvolupat en el departament d'IT de l'empresa SEAT S.A., on es dediquen a la producció i la fabricació d'automòbils. En aquest departament, una de les tasques que tracten són les incidències que tenen els treballadors. Emmagatzemen dades sobre la informació de la incidència en format de tiquets en una base de dades. Aquestes incidències són una part molt important en el rendiment de l'empresa, ja que tenen un gran impacte en la producció i en el treball de l'empresa.

Es vol realitzar una ajuda a l'empresa amb la gestió de la solució d'aquests tiquets d'incidències.

### 1.1 Objectius

L'objectiu principal del projecte és poder donar una solució o una aproximació a la solució a la incidència que té l'usuari, sent una capa intermèdia entre l'usuari i la solució, i així, poder evitar acudir a l'ajuda del suport de 5005 o minimitzar-la.

D'una banda, es vol arribar a una solució molt més ràpida perquè l'usuari segueixi amb el seu treball amb total normalitat. I, d'altra banda, es vol minimitzar costos en els contractes de l'ajuda oferta per 5005. Contra menys ajudes se sol·liciti, menys gent cal contractar.

El treball se centrarà en l'estudi de mètodes de Processament del Llenguatge Natural (NLP, per les seves sigles en anglès) per trobar bones solucions a les incidències, i després es plantejarà com implementar aquest estudi als programes i processos de l'empresa i que passi a ser una bona ajuda, ja que pot suposar un canvi econòmic.

### 1.2 Planificació

El projecte estarà organitzat en diferents fases a complir:

#### 1. ESTRATÈGIA

Com a primera fase, la primera cosa important a tenir en compte com a començament d'aquest projecte és entendre les dades amb els quals es tractarà i saber quina és la millor manera de tractar-los.

El conjunt de dades consta d'un número molt elevat de tiquets. Cada tiquet consta dels següents camps:

- Tiquet Type: si es tracta d'un incident o d'una sol·licitud.
- Tiquet ID: combinació única per a la identificació dels tiquets.
- CI: nom de l'aplicació on es troba l'incident.
- Title: títol que l'usuari li dona a l'incident.
- Description: descripció que l'usuari realitza en descriure el seu problema.
- Solution: es descriu la solució al problema de l'usuari.
- Open Time: data i hora la qual el tiquet ha estat creat.
- Close Time: data i hora la qual el tiquet ha estat solucionat.

- Current Estatus: informació sobre si s'ha donat solució al tiquet o no.
- Currents Assignment: grup al qual ha estat assignat aquest incident.
- Prio: camp numèric que indica la rapidesa en què la incidència s'ha de solucionar.
- Impact: camp numèric que indica l'impacte de la incidència dins del treball en SEAT.

Com a primera estratègia, es procedirà a realitzar subconjunts de dades. Primerament, es mira de separar les dades depenent de la seva CI, ja que pot ser que existeixin diferents aplicacions que tinguin mateix problema, però la solució a cada problema sigui totalment diferent. També es duran a terme filtrats per a tenir millors resultats. Es mirarà el nombre de tiquets de cada CI i la longitud del text a tractar, i a partir d'aquí, es podran ajuntar CI o mirar d'eliminar alguns tiquets que no serviran per a l'entrenament. A més a més, es mirarà la quantitat de tiquets que es generen a partir de cada CI, i s'estudiarà l'impacte que seria la solució d'aquests incidents en l'àmbit econòmic per a l'empresa, ja que l'objectiu principal és minimitzar costos.

Es crearan *dashboards* i gràfiques per a tenir una millor visualització d'aquests estudis. Cada subconjunt s'aplicarà al model o mètodes per separat.

#### 2. EVALUACIÓ

A partir dels subconjunts creats en la fase anterior, s'avaluarà si l'estudi realitzat és correcte o no. A partir d'aquesta anàlisi, es decidirà si s'ha de fer una nova estratègia i es tornarà a aquesta fase.

Aquesta avaluació consistirà a seleccionar un nombre petit de tiquets i aplicar diferents models o mètodes de NLP a partir dels camps *Title* i *Description* i s'ha de poder seleccionar el millor o els millors perquè pugui ser aplicat a totes les dades.

Es crearan *dashboards* i gràfiques per a tenir una millor visualització d'aquests estudis.

#### 3. APLICACIÓ

Una vegada completada les anteriors fases, ja tindriem les dades preparades i el mètode a utilitzar. Aquesta fase consisteix a posar en pràctica tot l'estudi realitzat en les anteriors fases i aplicar el mètode a totes les dades.

Podrà ser que trobem alguna errada a la pràctica amb totes les dades, així que es podria ajustar en qualsevol moment els paràmetres i les decisions preses en els estudis.

Treure resultats finals.

#### 4. CONCLUSIONS

Analitzar els resultats obtinguts i explicar conclusions en referència a la solució trobada i en el possible impacte a l'empresa. Donar una conclusió general del projecte, les dificultats trobades pel camí i les coses apreses.

Si el projecte es queda en un estudi per falta de temps o recursos, explicar la implementació a fer en un futur.

Si no, desenvolupar en una pròxima fase les implemen-  
tacions.

## 5. FASE OPTATIVA

App Service Desk SEAT és una aplicació creada pel departament on l'usuari emplena una sèrie de camps i crea un tiquet d'incidències. En aquest moment són enviats a 5005 per a la seva solució i és guardat a la base de dades.

Aquesta última fase consisteix a realitzar una nova interfície en l'aplicació amb els 3 tipus de propostes inicials. Parlar amb caps i persones creadores de l'aplicació per a piular codi i crear i connectar tot.

Implementacions per dur a terme: abans de donar-li una possible solució, es faran una sèrie de preguntes (arbre de preguntes) a l'usuari que ajudin el model a augmentar l'èxit (percentatge de probabilitat final) de trobar la solució, detallar un llistat d'accions per a ajudar a 5005 a acostar-se i arribar abans a la solució, crear el script que en executar-se solucioni el problema de l'usuari.

TAULA 1: TAULA DE PLANIFICACIÓ

Fases	Temps a dedicar
Estat de l'art	3 dies
Decisió de la millor estratègia amb ajuda de visualitzacions	5 setmanes
Aplicar a les dades i estudiar millor mètode	4 setmanes
Estudi dels resultats i fase obtativa	3 setmanes
Conclusions	2 setmanes
Revisió final	1 setmana

## 2 NLP

## 3 GESTIÓ DE LES DADES

## 4 MÈTODE A TRIAR

## 5 AVALUACIÓ DELS RESULTATS

## 6 CONCLUSIONS

## AGRAÏMENTS

## REFERÈNCIES

[1] <http://en.wikibooks.org/wiki/LaTeX>

[2] Referència 2

[3] Etc.

## APÈNDIX

### A.1 Secció d'Apèndix

... ..

..... ..

### A.2 Secció d'Apèndix

... ..