

ANÀLISI DE DADES

Març de 2023

Informe realitzat per:

Pere Girona Campi

 github.com/perot84

 pere.girona@onmail.com

 linkedin.com/in/peregirona



Anàlisi de dades ACAPPS

Aquest informe analitza i exposa els resultats del tractament de dades de extrets del CRM de l'entitat a 10 de març de 2023.

S'ha treballat a partir de tècniques de ciència de dades i de "Business Intelligence" per a analitzar les dades en tot el seu conjunt.

Objectius

Procediments

Resultats

Possibilitats de millora per part del servei

Annexes

Objectius

- Agilitzar l'entrada de dades
- Millorar l'entrada de dades
- Millorar el processament de les dades
- Facilitar la creació de justificacions i projectes
- Millorar la presa de decisions a partir de l'anàlisi i predicció de dades

Procediments

- Unió de totes les dades vinculades als usuaris, centrant-nos en les persones que es troben en cerca de feina.
- Predicció de quan triga una persona a trobar feina, des de que entra al servei.
- Detecció de millores en la recollida i gestió de les dades
- Creació d'un panell en Power BI on es visualitzen diferents dades

Procediment tècnic

Extracció de les dades

S'ha treballat a partir de les dades extretes de les següents base de dades

- Persones
- Relacions amb persones
- Seguiments
- Dades discapacitat
- Dades formatives
- Experiència laboral
- Candidatures

Neteja i creació de noves columnes

S'ha fet un procés de neteja de dades i creació de noves columnes entre les quals podem destacar:

- **Nivell formatiu** (s'han classificat els diferents nivells formatius recollits de cada persona mantenint el superior)
- **Quantitat de seguiments**
- **Temps a trobar feina**

Unió de la base de dades i creació de l'algoritme predictiu

S'ha creat una nova base de dades on queden recollides totes les dades pròpies de cada persona incloses les noves columnes creades i s'ha creat un algoritme per fer la predicció.

Resultats

- [Base de dades amb totes les dades pròpies de cada persona incloses les noves columnes creades](#)
- Predicció del temps que trigaria una persona trobarà feina des de que entra al servei.

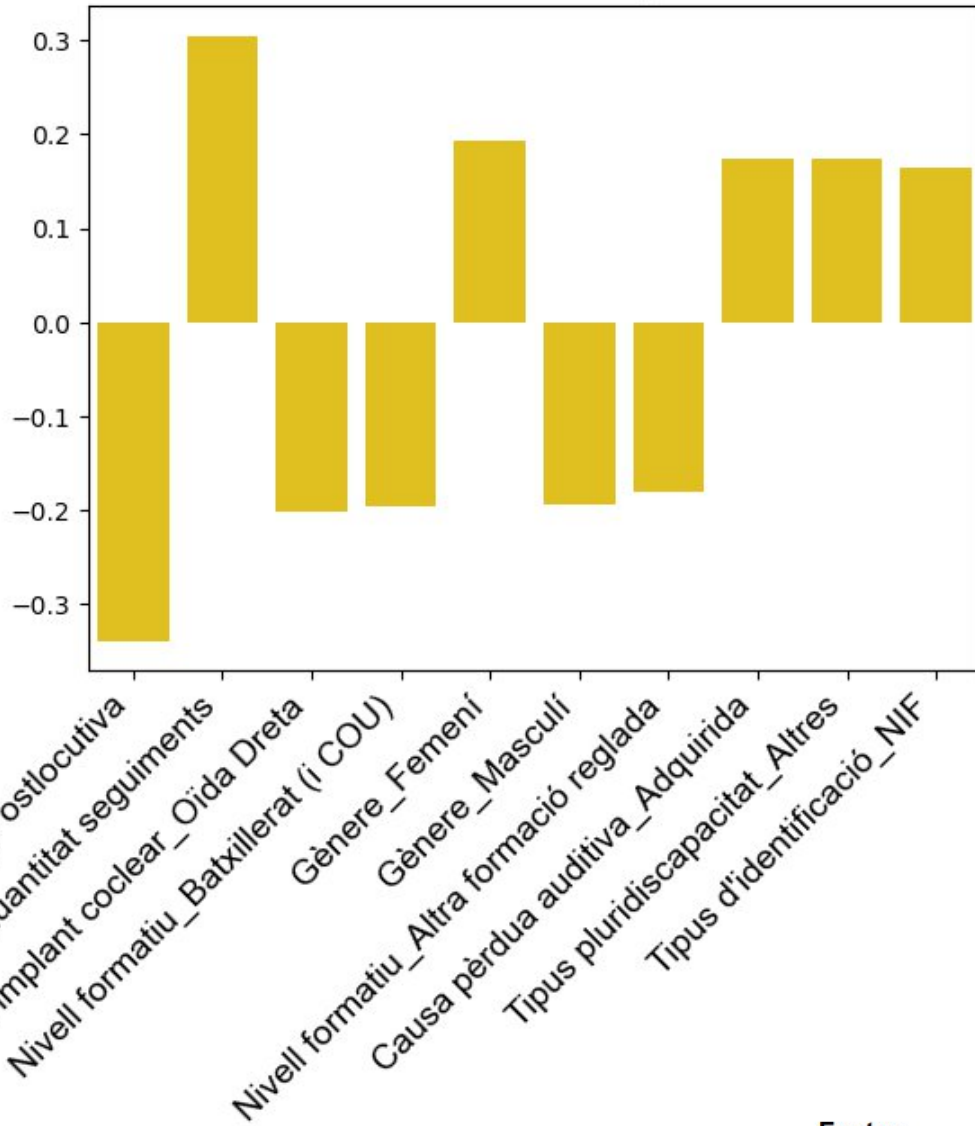
Després de la neteja de dades s'ha treballat amb una mostra de 77 persones amb un marge d'error de 102 dies.

Aquest resultat implica que el model té un error de 102 dies en la seva predicció. Cal destacar que hem treballat amb poques entrades (77) i eliminant algunes variables per no tenir suficients dades o ser de poca qualitat. Creiem que si es millora la qualitat i quantitat de dades es podrien millorar de manera significativa els resultats.

- [Informe de propostes de millora en la recollida i gestió de les dades](#)

- Anàlisi dels factors més determinants en el temps que triga una persona a trobar feina

Correlació dels factos en el temps de trobar feina



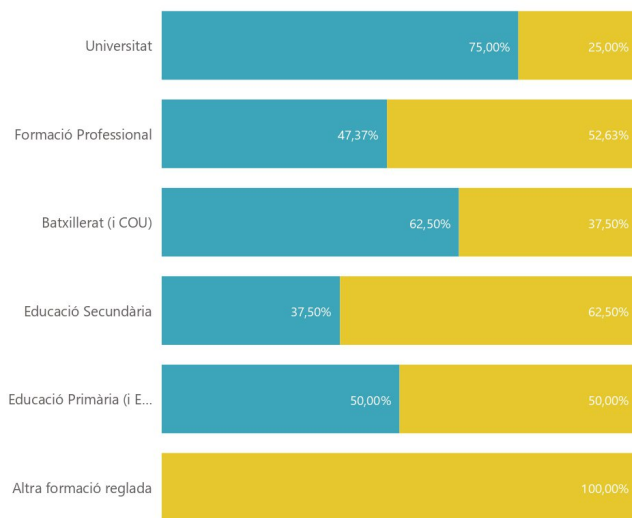
Factor	Valor
Moment apareix sordesa_Postlocutiva	-0.339201
Quantitat seguiments	0.304742
On implant coclear_Oïda Dreta	-0.201708
Nivell formatiu_Batxillerat (i COU)	-0.195988
Gènere_Femení	0.193359
Gènere_Masculí	-0.193359
Nivell formatiu_Altra formació reglada	-0.179889
Causa pèrdua auditiva_Acquirida	0.173007
Tipus pluridiscapacitat_Altres	0.172954
Tipus d'identificació_NIF	0.164330

- Anàlisi en clau de gènere de diferents aspectes de l'entitat
- [Arxiu Power BI](#)
- Exemples de possibles anàlisis

PARTICIPANTS SAES

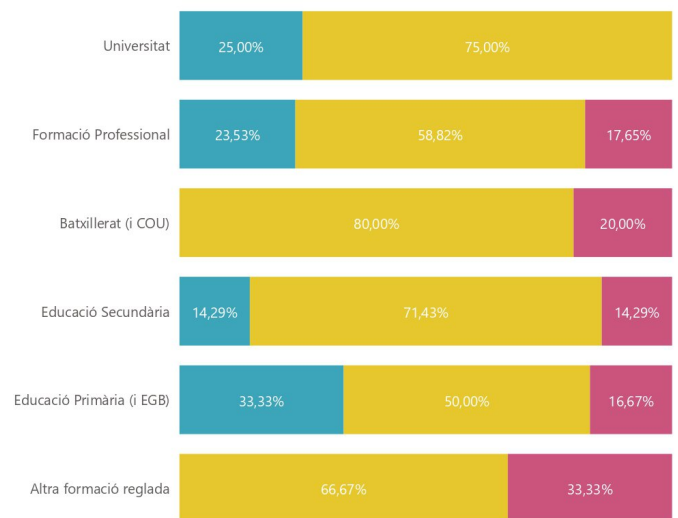
Gènere segons nivell formatiu

Femení Masculí



Sistema de comunicació segons nivell formatiu

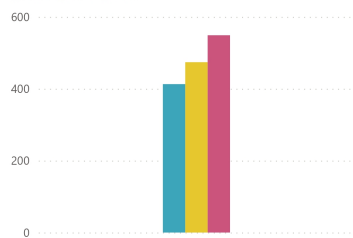
Bimodal Oral Signes



SEGUIMENTS

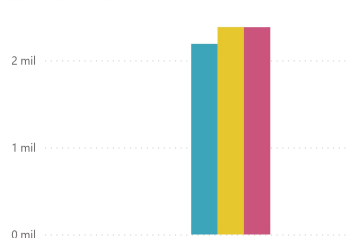
Nombre d'usuaris en seguiment

2020 2021 2022



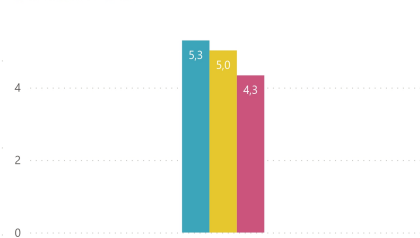
Quantitat de seguiments totals

2020 2021 2022



Mitjana de seguiments per persona

2020 2021 2022



Nombre d'usuaris en seguiment per gènere

2020 2021 2022



Possibilitats de millora per part del servei

- **Crear més columnes** (quantitat de seguiments per tipus...)
- **Fer altres prediccions**
 - A quantes ofertes s'ha d'inscriure una persona per a trobar feina
 - Evolució de la quantitat d'usuaris, d'empreses, de socis...
 - Predicció de subvencions
 - Factors que determinen l'aprovació o no de les subvencions
- **Automatitzar** la creació de les base de dades i de les prediccions a partir de la creació d'una API (protocol que connecta la base de dades del CRM amb el codi que crea les base de dades i les prediccions)

Annexes

- [Classificació catalana d'ocupacions 2011 \(ca\)](#)
- [Classificació catalana d'educació 2020 \(classificació del nivell de formació assolit\) \(ca\)](#)