

## Pràctica 4 - Arbres de decisió– Detecció de fatiga/estrès en jocs

Aion és un videojoc de rol en línia desenvolupat per NCSoft, llançat originalment el 2008 a Corea del Sud i després a altres regions. Està ambientat en el món fantàstic d'Atreia, que es va dividir en dues meitats arran d'un cataclisme: Elysea, habitada pels Elyos, i Asmodae, pels Asmodians. Aquestes dues faccions jugables viuen en constant conflicte, mentre



els Balaur, una raça de dracs poderosos i caòtics, representen una amenaça comuna. Els jugadors poden crear i personalitzar personatges amb diverses classes, com guerrers, mags o sacerdots, i anar desenvolupant habilitats a mesura que progressen en el joc. Un dels aspectes més destacats d'Aion és el seu sistema de vol, que permet als jugadors planar o lluitar en l'aire, aportant una dimensió estratègica tant en combats PvE (jugador contra entorn) com PvP (jugador contra jugador). El joc ofereix missions, exploració, un sistema de gremis i batalles a gran escala en zones com l'Abyss, un territori neutral en disputa. Tot i el seu èxit inicial, Aion ha evolucionat cap a un model free-to-play amb actualitzacions regulars per mantenir l'interès dels jugadors, i destaca pel seu estil visual detallat i una mitologia profunda, encara que la seva popularitat ha disminuït en comparació amb altres jocs.

Com a membres de l'equip de biomedicina de NCSoft, l'empresa desenvolupadora d'Aion, se us ha encarregat una tasca crítica: analitzar les dades del joc per detectar possibles signes de fatiga o estrès en els jugadors reals (no en els personatges). Aquest projecte pel qual se us ha contractat sorgeix arran de les noves regulacions internacionals sobre benestar digital, que obliguen les empreses de videojocs a monitoritzar i protegir la salut mental i física dels usuaris, evitant problemes com l'addicció al joc o l'esgotament per sessions de joc prolongades. La detecció d'aquests estats ens permetrà implementar mesures preventives, com avisos personalitzats o límits de temps de joc, per complir amb les normatives i millorar l'experiència dels nostres usuaris, alhora que protegim la reputació de l'empresa i evitem possibles sancions legals.

Com que analitzar manualment el comportament de milers de jugadors és una tasca repetitiva i costosa que han de fer metges i en contacte directe amb els jugadors. S'ha fet només l'estudi sobre un petit conjunt dels jugadors. Disposem doncs del dataset "dades\_aion.csv", que conté dades del joc AION i l'avaluació dels més del 9000 comptes analitzats amb la informació de si els jugadors mostren o no signes de fatiga o estrès. Les variables que ens arriben han estat proporcionades pel departament de dades i, malauradament, no en tenim més descripció que el nom de la variable.

### Tasques a realitzar:

1. **Preparació del dataset (1 punt):** Prepareu el dataset per construir correctament els models (no elimineu variables). Justifiqueu cadascun dels passos i decisions que preneu, explicant com heu preparat les dades.

2. **Construcció d'un arbre de decisió inicial (1 punt):** Creeu un arbre de decisió utilitzant tots els paràmetres per defecte, excepte el paràmetre `criterion="entropy"`, per predir si un jugador pateix estres. Compareu les mètriques obtingudes amb el conjunt d'entrenament i el de test. Calculeu la taxa de falsos positius i la taxa de falsos negatius. És un bon model? Justifiqueu la resposta.
3. **Modificació de l'arbre de decisió (1 punt):** Construïu un nou arbre de decisió, però aquest cop afegiu el paràmetre `max_depth=5`. Observeu si hi ha diferències en les puntuacions obtingudes amb els conjunts d'entrenament i de test en comparació amb els resultats anteriors. Expliqueu qualsevol canvi que detecteu i les possibles raons d'aquestes diferències.
4. **Comparació amb els algorismes Dummy (2 punt)** Utilitzeu un algorisme "dummy" amb estratègia "most\_frequent" i compareu resultats amb els obtinguts amb l'arbre de decisió dels apartats anteriors. Quin problema / problemes veieu en aquest datasset? Quin dels models fets fins ara creieu que es el que funciona millor?
5. **Optimització del model (5 punts):** Milloreu el model (no es pot canviar l'algorisme, es una pràctica d'arbres de decisió). Justifiqueu detalladament cada canvi i pas que realitzeu, explicant per què heu pres aquestes decisions (per exemple, ajustament de paràmetres, preprocessament de dades, etc.). Es valorarà més la correcció de les justificacions i els procediments que la simple obtenció de bons resultats. Finalment, apliqueu el model a tots els usuaris que no s'havia analitzat manualment.

**Lliurament:** Com que els resultats poden variar entre execucions a causa de l'aleatorietat, heu de lliurar el codi utilitzat per respondre les preguntes i els resultats obtinguts a mes de les conclusions sol·licitades a cada apartat.