

METODI ED ALGORITMI

Domanda 1

- Gli appelli con stessi CDS, AD e data vengono aggregati. Consideriamo, quindi, un unico appello l'eventuale divisione per docente
- Dati calcolati (output):
 - Numero di iscrizioni
 - Numero di appelli
 - Numero medio di iscrizioni per ciascun appello
- Raggruppamento per CDS e per AD. Teniamo anche i dati delle distribuzioni del numero di iscritti per singolo appello (in vista della visualizzazione)

Domanda 2

- Calcolo della difficoltà del singolo appello
 - Numero di iscritti
 - Tasso di superamento (calcolato come il rapporto tra numero di promossi e numero di iscritti)
 - Voto aggregato (calcolato mediando voto medio e voto mediano)
 - Voto medio NOT NULL and voto mediano NOT NULL
 - $ts \geq 0.5$
 - $\text{voto agg} = (\text{voto medio} + \text{voto mediano}) / 2$
 - Voto medio NOT NULL and voto mediano NULL
 - $ts > 0$ and $ts < 0.5$
 - Intervallo (con approx a 3 cifre) $[0.001, 0.499]$
 - Ampiezza dell'intervallo è 0.498
 - Corrisponde ai voti in trentesimi $[1, 17]$
 - Dividiamo 0.498 in 16 parti uguali, ottenendo l'ampiezza del singolo intervallo all'interno del quale cadono i voti
 - Per ottenere il voto finale: $(\text{rapporto tra } ts \text{ ed ampiezza del singolo intervallo}) + 1$
 - Voto medio NULL and voto mediano NULL $\rightarrow 0$
 - Per ciascun AD \rightarrow parametri utilizzati per calcolare la classifica finale di difficoltà
 - Media tentativi per ciascun studente
 - Mediana del tasso di superamento
 - Mediana del voto aggregato dei singoli appelli
 - **Calcolo del ranking finale per AD**
 - Calcolo della classifica per ciascun parametro all'interno di ciascun CDS
 - Tasso superamento mediano e voto aggregato mediano (ASC)
 - Media tentativi (DESC)
 - Somma pesata dei rank dei singoli parametri per ottenere il ranking finale
 - $TS (0.4) + \text{voto agg} (0.3) + \text{media tentativi} (0.3)$
- ★ Vengono considerati solamente gli appelli chiusi!

Domanda 3

- Per ciascuna data
 - Numero di appelli di AD distinte per ciascun CDS (commitment count)
- Per ciascun CDS
 - Numero di appelli (totale)
 - Corrisponde alla somma del numero di appelli di AD distinti per data
 - Numero di AD
 - Numero di date con almeno un overlap di AD (numero_date_overlap)
 - Numero di date (con almeno un appello)
 - Numero di appelli in date con almeno un overlap (numero_esami_date_overlap)
- **Calcolo del commitment rate**
 - Params considerati
 - Calcolo della percentuale (ratio) di date con overlap sul totale di date (A)
 - Calcolo della media del numero di appelli in date con almeno un overlap (B)
 - Calcolo del rapporto tra numero di AD e numero di appelli (C)
 - $\text{Commitment Rate} = 0.7 * A * B + 0.3 * C$
- **Calcolo del percentile rank**, idea: qual è la percentuale (ratio) di CDS con commitment rate minore

Domanda 4

- Per ogni AD
 - Numero di appelli
 - Numero di iscrizioni medio (a ciascun appello)
 - Voto medio
 - Voto medio dell'appello: calcolato utilizzando il voto medio iniziale pesato per il numero di promossi
 - Voto medio AD: somma dei voti medi degli appelli diviso la somma del numero di promossi
 - Voto medio normalizzato (min-max, range 18 - 30)
 - Rapporto tra promossi ed iscritti
- Thresholds (campione rilevante per il nostro algoritmo)
 - Teniamo solamente le AD con almeno 3 appelli
 - In media devono esserci almeno 2 iscritti per appello (per AD)
- **Rating finale**
 - Somma pesata: voto medio normalizzato (0.7) + rapporto tra promossi ed iscritti (0.3)
 - Ordinamento per CDS sulla base del rating ed, a pari merito, considerando il numero di appelli (ragionamento, più appelli significa più rilevante)
- ★ Raggruppiamo per CDS ed AD, non facendo distinzione per docente!
- ★ Abbiamo considerato solamente gli appelli chiusi!

Domanda 5

1. Consideriamo molteplici parametri per ottenere il **F&F index** per ciascuno studente all'interno del proprio CDS
 - **Parametri Fast**
 - *Periodo di attività* (calcolato come mesi che intercorrono tra il primo e l'ultimo appello sostenuto)
 - **Parametri Furious**
 - *Voto medio* (media dei voti degli esami superati)
 - *Numero di esami superati*
 - *Tasso di successo* (calcolato come il rapporto tra numero di esami superati ed il numero di iscrizioni)
 2. **Normalizzazione dei parametri (min-max normalization)**, con lo scopo di averli tutti nella stessa scala (ovvero nel range $[0,1]$) e di poter quindi computare l'indice finale
 - *Voto medio normalizzato* $\rightarrow [0, 30]$
 - *Numero di esami superati* $\rightarrow [0, \text{massimo numero di esami superati nel CDS}]$
 - *Tasso di successo* (già compreso nel range)
 - *Periodo di attività* $\rightarrow [1, 13]$
 - L'ultimo appello chiuso è nel mese di novembre!
 3. Calcolo del **F&F index** attraverso una somma pesata dove ogni parametro ha peso 0.25
 - Attenzione mesi di attività (meno mesi significa più fast): $(1 - \text{mesi_att}) * 0.25$
- ★ Vengono considerati solamente gli appelli chiusi!

Domanda 6

- Per ciascuno studente, AD e CDS
 - *Numero insufficienze*
 - *Numero di ritiri*
 - *Numero di assenze*
 - **Trial&Error index** (per singolo studente all'interno di un AD)
 - *Somma pesata delle bocciature*: insufficienze (1.5), ritiri (1) ed assenze (0.5)
 - **Trial&Error AD**: media T&E degli studenti di quello specifico AD
 - **Trial&Error CDS**: media T&E degli AD di quello specifico CDS
- ★ Vengono considerati solamente gli appelli chiusi!

Domanda 7 - Gold Mortarboard

Idea: algoritmo per definire i migliori studenti dell'anno, meritevoli di una borsa di studio

- Prendere i migliori studenti per ciascun CDS
- ***Fairness nell'assegnazione***
 - Il numero di studenti selezionati per CDS sarà proporzionale al numero di "studenti attivi" di quello specifico CDS
 - Assegnare le borse di studio separatamente per L2, LM ed LM5 + LM6 (tre categorie)

Algoritmo

- Per ciascuno studente (all'interno di un CDS) vengono considerati i seguenti parametri
 - *Fast and Furious Index*
 - *Average Trial and Error Index* (parametro relativo alla difficoltà degli AD superati)
 - Media del T&E index (per AD) di tutte le AD che lo studente ha superato
 - *Average Difficulty Index* (parametro relativo alla difficoltà degli AD superati)
 - Consideriamo il rank pesato di difficoltà normalizzato (min-max within the CDS) per AD (ottenuto dalla query 2) e facciamo la media per tutte le AD che lo studente ha superato
- **Calcolo del Gold Mortarboard Score**
 - *Aggregazione dei parametri* descritti in precedenza: media
 - Threshold: numero minimo di esami superati
- Decidere quanti studenti selezionare considerando anche la fairness