



Notice du micro service de notation

Auteur :
Guillaume DE BIGAULT DE GRANRUT

8 juillet 2020

Table des matières

I	Les résultats calculés	2
I.1	Traitement de la jockerisation	2
I.2	Pour un critère	3
I.2.1	Les résultats individuels	3
I.2.2	Les résultats globaux (moyenne sur le critère)	4
I.3	Pour une phase ou une activité	5
I.3.1	Les résultats individuels	5
I.3.2	Les résultats globaux	5
II	Les paramètres en entrée	6
II.1	Pour les critères	6
II.2	Pour les phases et activités	7
III	Les données en sortie	8
IV	architecture de l'appli	9
IV.1	Le module de calcul par critère	10
IV.2	Le module de calcul par phase ou activité	11
V	Remarques liées à l'utilisation par l'appli Serpico	11
VI	Mise en prod et utilisation	11

INTRODUCTION

Ce micro service permet de calculer tous les résultats nécessaires au fonctionnement de la solution proposée par Serpico. Ce service, totalement indépendant du reste de la solution, prends en entrée un json avec les données nécessaires, et renvoie en sortie un autre json contenant les résultats calculés. Ce service permet la jockerisation. Ce service est développé en python.

I. - LES RÉSULTATS CALCULÉS

Cette appli calcule des résultats pour des critères, des phases (stage), et des activités. Il y a deux types de résultats :

- Les résultats individuels : ce sont les résultats réalisés par une entité, que ce soit un user, une team, ou même une phase. Ils se décompose en deux types de valeurs :
 - Les résultats weighted : les résultats pondérés par le poids des notés et des notants, selon utilisation dans les formules.
 - Les résultats equal : les résultats non pondérés, ce qui équivaut à utiliser une pondération de 1.

Dans un souci de simplification, les formules seront détaillées une seule fois, avec la pondération.

- Les résultats globaux : c'est des moyennes sur les différents critères, ainsi que quelque autres paramètres calculés. Ils se décomposent en 4 cas (plus en comptant les valeurs avec un maximum théorique) :
 - Les valeurs sur les user : moyennes calculées sur les user, et non les team. Elles se décomposent également en weighted et equal.
 - Les valeurs sur les team : moyennes calculées sur les team, et pas les user. Elles se décomposent également en weighted et equal.

Dans un souci de simplification et de lisibilité, les formules seront détaillées que dans un seul cas (pondéré). Il suffira de remplacer par les populations correspondantes.

1. Traitement de la jockerisation

Le moyen le plus simple de traiter la jockerisation est simplement de ne rien renseigner quand une note est manquante, l'application bouclera sur les notes renseignées, et fera les calculs de pondérations en conséquence. Le seul défaut est de devoir calculer la somme des poids pour chaque entité, ce qui est peu couteux. Pour le calcul des résultats globaux, on utilise les moyennes comme avant, car changer les formules aurait peu de sens, et apporterait uniquement de la complexité pour rien, et ne causerait que des problèmes.

2. Pour un critère

Cette partie est la plus importante. Elle process l'ensemble des notes. Voici les résultats calculés :

a) Les résultats individuels

On note :

- $\text{grade}[\text{grader} \rightarrow \text{graded}]$: la note donnée par grader à graded
- $\text{graders}(u)$: personnes évaluant u (si u appartient à une team, prends en compte la note attribuée à sa team)
- $\text{graded}(u)$: les personnes évaluées par u (peut varier à cause de la jockerisation)
- graders : ensemble des notants du critère (décrit soit l'ensemble des user, soit des team contenant au moins un user notant).
- graded : ensemble des entités évaluées pour ce critère.
- $\text{result}(u)$: le résultat de u
- $\text{weight}(u)$: le poids de u
- $\text{relativeResult}(u)$: le résultat relatif de u
- $\text{stdDev}(u)$: l'écart type de u
- team : membres de la team de

Les résultats individuels sont :

- **Result** : simplement la moyenne de résultat pour un critère, résultat concernant les notés, soit les actifs et les passifs.

$$\text{result}(u) = \sum_{g \in \text{graders}(u)} \frac{\text{grade}[g \rightarrow u] \times \text{weight}(g)}{\sum_{g \in \text{graders}(u)} \text{weight}(g)}$$

- **RelativeResult** : formule appliquée à la moyenne de résultat, qui donne une idée de celle-ci, sans avoir les bornes de la note.

$$\text{relativeResult}(u) = \frac{\text{result}(u) - \text{lb}}{\text{ub} - \text{lb}}$$

- **stdDev** : concerne les notants (les tiers utilisateur et les actifs). Cet écart type est réalisé sur les notes données et non pas reçues. Il décrit à quel point l'utilisateur est d'accord avec les autres notants. Il vaut :
 - pour les user

$$\text{stdDev}(u) = \sqrt{\sum_{g \in \text{graded}(u)} \frac{(\text{grade}[u \rightarrow g] - \text{result}(g))^2 \times \text{weight}(g)}{\sum_{g \in \text{graded}(u)} \text{weight}(g)}}$$

- pour les team

$$\text{stdDev}(\text{team}) = \sum_{g \in \text{team}} \frac{\text{stdDev}(g) \times \text{weight}(g)}{\sum_{g \in \text{team}} \text{weight}(g)}$$

- **devRatio** : un écart type ne parlant pas forcément, Serpico propose un ratio entre l'écart type et un maximum théorique d'écart type. Il vaut :

$$\text{devRatio}(u) = \frac{\text{stdDev}(u)}{\text{maxStdDev}}$$

maxStdDev est l'écart type maximum théorique possible. Son calcul est détaillé plus bas. Il prend une valeur différente selon qu'on traite un user ou une team.

b) Les résultats globaux (moyenne sur le critère)

Les résultats globaux sont (pour rappel, en 4 fois, selon pondération et team ou user) :

- averageResult : moyenne des résultats sur le critère. Cette valeur n'est calculée que si les bornes de notation sont les mêmes pour tous les critères (n'a pas de sens sinon, est set à `null`).

$$\text{averageResult} = \sum_{g \in \text{graded}} \frac{\text{result}(g) * \text{weight}(g)}{\sum_{g \in \text{graded}(u)} \text{weight}(g)}$$

- averageRelativeResult : même principe que pour le résultat relatif individuel.

$$\text{averageRelativeResult} = \frac{\text{averageResult} - \text{lb}}{\text{ub} - \text{lb}}$$

- averageStdDev : moyenne des écarts types.

$$\text{averageStdDev} = \sum_{g \in \text{grader}} \frac{\text{stdDev}(g) \times \text{weight}(g)}{\sum_{g \in \text{grader}} \text{weight}(g)}$$

- maxStdDev : maximum théorique des écarts types.

$$\text{maxStdDev} = \sum_{g \in \text{graded}} \frac{\max(\text{ub} - \text{result}(g), \text{result}(g) - \text{lb})}{\sum_{g \in \text{graded}} \text{weight}(g)}$$

- inertia : inertie des écarts type :

$$\text{inertia} = \sum_{g \in \text{grader}} \frac{\text{stdDev}(g)^2 \times \text{weight}(g)}{\sum_{g \in \text{grader}} \text{weight}(g)}$$

- maxInertia : carré de maxStdDev

$$\text{maxInertia} = \text{maxStdDev}^2$$

- les devRatio : le ratio de l'Inertie

$$\text{devRatio} = \frac{\text{inertia}}{\text{maxInertia}}$$

3. Pour une phase ou une activité

Serpico propose de calculer un certains nombre de moyennes sur les phases et activités. Comme le principe est le même pour une phase ou une activité, nous ne détaillerons les calculs qu'une seule fois, il faut prendre la population parmi les critères ou les phases, selon qu'on fasse les moyennes pour un phase ou une activité. Encore une fois, ces résultats sont calculés dans le cas pondéré et equal, et user ou team pour les valeurs globales.

On utilise les notations suivantes, dans le cas où elles sont définies :

- $\text{result}[u, \text{crit}]$: le résultat (moyenne sur un critère ou phase) de u sur crit (un critère ou une phase).
- $\text{relativeResult}[u, \text{crit}]$: le résultat relatif de u sur crit .
- $\text{stdDev}[u, \text{crit}]$: l'écart type (moyenne sur un critère ou phase) de u sur crit (un critère ou une phase).
- $\text{devRatio}[u, \text{crit}]$: le ratio de déviation de u sur crit .
- $\text{weight}(\text{crit})$: le poids du critère, nécessairement défini.
- Crit : soit l'ensemble des critères, ou des phases, selon l'élément traité.

a) Les résultats individuels

- les moyennes de résultats (on parle aussi de résultat absolu dans le code existant de Serpico).

$$\text{averageResult}(u) = \sum_{c \in \text{crit}} \frac{\text{result}(u)[u, c]}{\sum_{c \in \text{crit}} \text{weight}(c)}$$

- les moyennes de résultats relatifs

$$\text{averageRelativeResult}(u) = \sum_{c \in \text{crit}} \frac{\text{relativeResult}[u, c]}{\sum_{c \in \text{crit}} \text{weight}(c)}$$

- l'écart type moyen

$$\text{averageStdDev} = \sum_{c \in \text{crit}} \frac{\text{stdDev}[u, c]}{\sum_{c \in \text{crit}} \text{weight}(c)}$$

- le ratio de déviation moyen :

$$\text{averageRatio} = \sum_{c \in \text{crit}} \frac{\text{devRatio}[u, c]}{\sum_{c \in \text{crit}} \text{weight}(c)}$$

b) Les résultats globaux

On fait la moyenne de la même manière, sur tous les paramètres globaux calculés précédemment. Je n'ai pas détaillé les calculs, car c'est vraiment la même chose qu'au dessus. Si tu veux, je le fais, et sinon ça sera laissé au lecteur.

II. - LES PARAMÈTRES EN ENTRÉE

Ce micro service prend en paramètre un json qui contient toutes les informations nécessaires pour calculer tous les résultats détaillés précédemment. Voici ci dessous les paramètres nécessaires :

1. Pour les critères

```

{
  "userWeights": {
    "id_user": value,
  },
  "teams": {
    "id_team": [
      "id_participant",
    ]
  }
  "teamWeight": {
    "id_team": value,
  }
  // Not yet implemented
  "legitimacy": {
    "id_graded": {
      "grader": value,
    }
  }

  "criterias": {
    "criteria_id": {
      "lowerbound": value,
      "upperbound": value,
      "user_grades": {
        "grader_id": {
          "graded_id": value // note du graded, donnee par le grader
          "graded_id2" : value
        }
      }
      "team_grades" {
        "grader_id": {
          "team_id": value // note de la team, donnee par le grader (un
            user)
          "team_id2" : value
        }
      }
    }
  }
}

```

est ce que c'est assez clair, ou est ce que je dois détailler plus ?

2. Poue les phases et activités

```

{
  "globalData":{
    "weight" :
    "averageUsersWeightedResult" : value,
    "averageUserslEqualResult." : value,
    "averageTeamWeightedResult" : value,
    "averageTeamEqualResult" : value,
    //RelativeResult
    "averageUsersWeightedRelativeResult" : value,
    "averageUserslEqualRelativeResult" : value,
    "averageTeamWeightedRelativeResult" : value,
    "averageTeamEqualRelativeResult" : value,
    //StdDev
    "averageUsersWeightedStdDev" : value,
    "averageUsersEqualStdDev" : value,
    "averageTeamWeightedStdDev" : value,
    "averageTeamEqualStdDev" : value,
    //MaxStdDev
    "maxUsersWeightedStdDev" : value,
    "maxUsersEqualStdDev" : value,
    "maxTeamWeightedStdDev" : value,
    "maxTeamEqualStdDev" : value,
    //Inertie
    "usersWeightedInertia" : value,
    "usersEqualInertia" : value,
    "teamWeightedInertia" : value,
    "teamEqualInertia" : value,
    //MaxInertie
    "maxUsersWeightedInertia" : value,
    "maxUsersEqualInertia" : value,
    "maxTeamWeightedInertia" : value,
    "maxTeamEqualInertia" : value,
    //Ratio
    "usersWeightedRatio" : value,
    "usersEqualRatio" : value,
    "teamWeightedRatio": value,
    "teamEqualRatio" : value,

    "user" :{
      "id_participant":{
        "weightedResult": value,
        "equalResult":value,
        "weightedRelativeResult":value,
        "equalRelativeResult":value,
        "weightedStdDev":value,
        "equalStdDev":value,
        "weightedRatioStdDev": value,
        "equalRatioStdDev": value
      }
    }
    "teams":{
      "id_participant":{
        "weightedResult": value,
        "equalResult": value,

```



```

        "weightedRelativeResult": value,
        "equalRelativeResult": value,
        "weightedStdDev": value,
        "equalStdDev": value,
        "weightedRatioStdDev": value,
        "equalRatioStdDev": value
    }
}
}
}

```

est ce que c'est assez clair, ou est ce que je dois détailler plus ?

III. - LES DONNÉES EN SORTIE

Les données en sortie auront toujours la même structure, que ce soit pour des calculs sur un critère, une phase, ou une activité. Ça permet notamment de faire une seule fonction de traitement dans l'appli principale.

```

{
  "id": {

    "averageUsersWeightedResult" : ,
    "averageUsersEqualResult" : ,
    "averageTeamWeightedResult" : ,
    "averageTeamEqualResult" : ,
    //RelativeResult
    "averageUsersWeightedRelativeResult" : ,
    "averageUsersEqualRelativeResult" : ,
    "averageTeamWeightedRelativeResult" : ,
    "averageTeamEqualRelativeResult" : ,
    //StdDev
    "averageUsersWeightedStdDev" : ,
    "averageUsersEqualStdDev" : ,
    "averageTeamWeightedStdDev" : ,
    "averageTeamEqualStdDev" : ,
    //MaxStdDev
    "maxUsersWeightedStdDev" : ,
    "maxUsersEqualStdDev" : ,
    "maxTeamWeightedStdDev" : ,
    "maxTeamEqualStdDev" : ,
    //Inertie
    "usersWeightedInertia" : ,
    "usersEqualInertia" : ,
    "teamWeightedInertia" : ,
    "teamEqualInertia" : ,
    //MaxInertie
    "maxUsersWeightedInertia" : ,
    "maxUsersEqualInertia" : ,
    "maxTeamWeightedInertia" : ,
    "maxTeamEqualInertia" : ,
    //Ratio
    "usersWeightedRatio" : ,
    "usersEqualRatio" : ,
    "teamWeightedRatio": ,

```

```

    "teamEqualRatio" : ,

    "user" :{
      "id_participant":{
        "weightedResult":,
        "equalResult":,
        "weightedRelativeResult":,
        "equalRelativeResult":,
        "weightedStdDev":,
        "equalStdDev":,
        "weightedRatioStdDev":,
        "equalRatioStdDev":
      }
    }

    "team" :{
      "id_participant":{
        "weightedResult":,
        "equalResult":,
        "weightedRelativeResult":,
        "equalRelativeResult":,
        "weightedStdDev":,
        "equalStdDev":,
        "weightedRatioStdDev":,
        "equalRatioStdDev":
      }
    }
  }
}

```

Remarque sur "id". Cette valeur n'a de sens que pour les critères, car ce framework peut en traiter plusieurs en même temps. Dans le cas des stage et des activity, cette valeur vaut "step". (Ne pas la changer sans mettre à jour l'appli principale).

Le choix est fait pour l'instant de ne pas regrouper les team et user, la principale raison qui a motivé ce choix est la gestion des clés. En effet, un user et une team peuvent avoir la même clé, en les séparant ainsi, on peut les renseigner avec leur vraie clé, ce qui simplifiera l'utilisation du micro service par le reste de l'appli Serpico.

IV. - ARCHITECTURE DE L'APPLI

On propose une solution orientée objet, et non pas un calcul matriciel, pour plusieurs raisons.

- Une telle matrice est creuse, et c'est pire avec la jockerisation, ce qui implique de la perte en mémoire, et en performance. Surtout, ça perd en lisibilité pour le code.
- En revanche, ce choix permet de coller à la structure de donnée, et de gagner en lisibilité, et de modulariser les calculs beaucoup plus facilement. Si jamais ils devaient être modifiés dans le futur, il suffira de juste modifier certaines fonctions, et pas tout le calcul matriciel.

Le code se décompose en deux parties. La première, pour le calcul par critère, et la deuxième pour le calcul par phase/activité.

1. Le module de calcul par critère

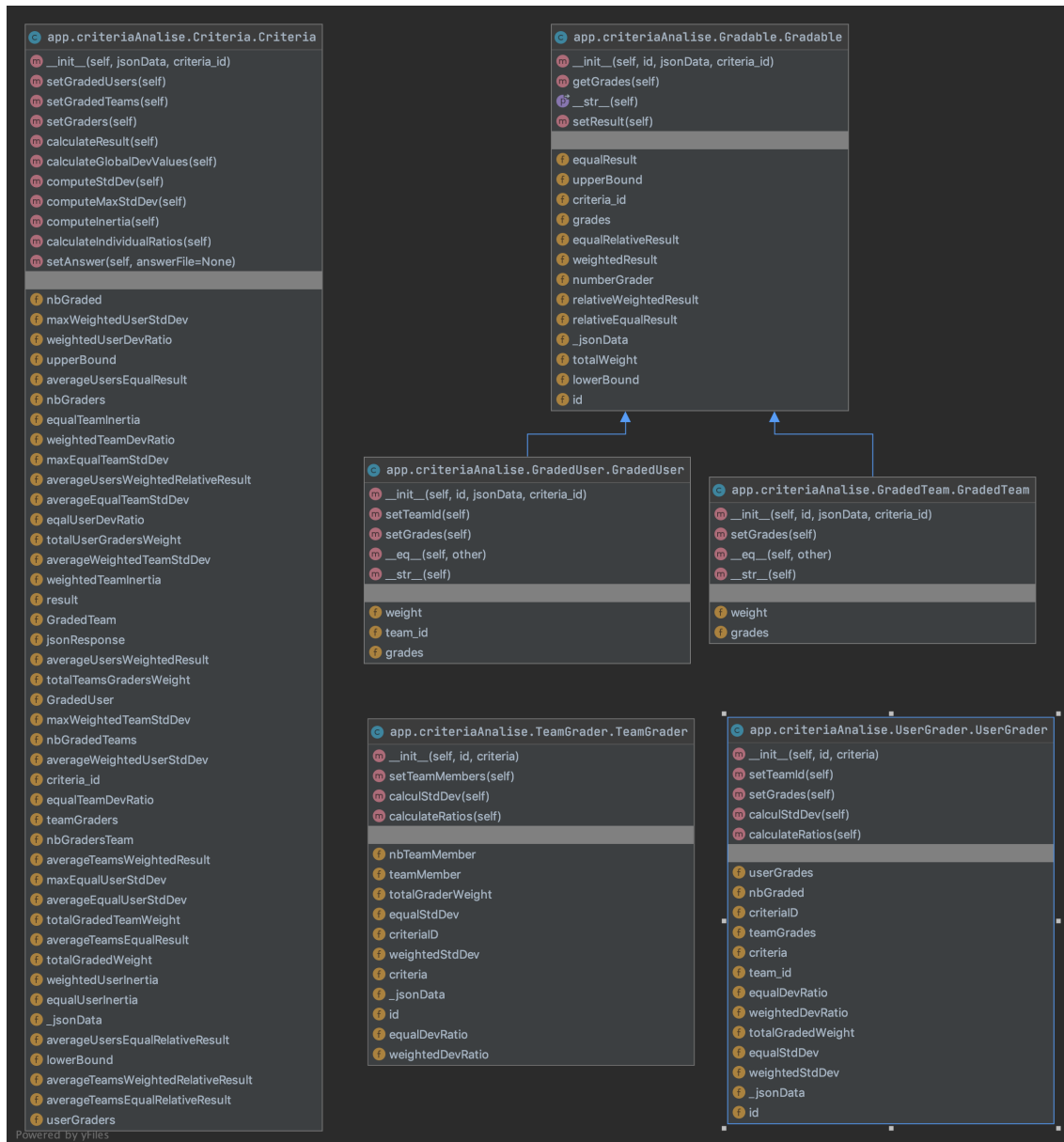


FIGURE 1 – Diagramme de classe de la partie calcul par critère

2. Le module de calcul par phase ou activité

Cette partie a été finalement faite en fonctionnel et non pas en objet, car ce n'était pas nécessaire au vu de la simplicité de cette partie.

V. - REMARQUES LIÉES À L'UTILISATION PAR L'APPLI SERPICO

Il a été décidé de séparer complètement les résultats liés à l'évaluation de phase et de user. Ainsi, ce service est appelé deux fois dans le cas où une phase est évaluée et la phase et les user. Dans le cas où on appelle le framework quand la phase est évaluée, elle est considérée comme un user passif, avec un poids de 1, et un id de -1 (le poids n'importe peu, dans ce cas la phase est le seul utilisateur noté).

VI. - MISE EN PROD ET UTILISATION

Le code est déployé sur le premier serveur de Serpico, voir le Readme pour plus de détails.