

Evaluation des dossiers sur les mathématiques du Big Data

Les « 12 travaux d'AstéRix »

Daif Hakim

27/01/2021

Modèles de régression

1. Critères d'évaluation

Dans ce dossier, nous allons parcourir et faire une évaluation des travaux de mes camarades en mathématiques du Big Data:

Nos critères seront :

- 1. Visuel et l'organisation du dossier:** Es ce que la mise en page est propre, agréable à lire ou non ? Le travail est-il organisé (chapitres, sections, sous-sections, Parties)
- 2. Qualité du latex et des formules mathématiques :** Es ce que les formules sont claires ou pas ? Es ce que le dossier contient le code Latex ou le rapport sous format Rmd ?
- 3. Compréhension de l'idée générale:** Les Auteurs arrivent-ils à transmettre l'idée générale ?
- 4. Explication et compréhension des formules mathématique et des concepts fondamentaux:** Es ce que les formules et les notions sont bien expliquées ?
- 5. Difficulté et originalité du thème :** Le sujet abordé est-il original/difficile ou pas ?

2. Lien vers le document commenté

Le document évalué dans le cadre de ce rendu a été fait par Nina Zoumanigui sur Modèles de régression .

lien vers leur Github: <https://github.com/Nina809/PSBX/blob/main/Regression.Rmd>

le dossier en question est **Regression.Rmd**

3. Synthèse de la présentation

Les modèles de régression sont construits dans le but d'expliquer (ou prédire, selon la perspective de l'analyse) la variance d'un phénomène (variable dépendante) à l'aide d'une combinaison de facteurs explicatifs (variables indépendantes).

La régression linéaire est appelée multiple lorsque le modèle est composé d'au moins deux variables indépendantes. À l'inverse, un modèle de régression linéaire simple ne contient qu'une seule variable indépendante.

le travail de Nina Zoumanigui est subdivisé en 3 partie : Introduction (une idée général sur le sujet illustré par la citation de Émile Durkheim), une partie consacrée à la regression simple une autre et une dernier pour la regression multiple.

Explication des formule

Pour la partie régression simple l'auteur ne nous a pas fourni malheureusement l'aspect mathématique de ce type de régression, il a seulement tenté de l'expliquer d'une façon théorique, néanmoins on trouve un exemple d'application sur R, dans ce cadre, Nina Zoumanigui a pris une base de données puis il a défini deux variables (une explicative et l'autre à expliquer) puis il construit le modèle de régression simple

Dans la deuxième partie "régression multiple", l'auteur n'a pas fourni assez d'éléments pour expliquer la régression multiple (ni application sur R ni aspect mathématiques), il s'est contenté juste de donner le modèle de la régression multiple sans explication

$$y_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip} + e_i, i = 1, \dots, n$$

$$y_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip} + e_i, i = 1, \dots, n$$

l'auteur n'a pas pu reproduire le modèle de la régression multiple à travers le code latex.

5. Evaluation

*1. Visuel et l'organisation du dossier: ** Le visuel est pas mal, on peut dire que le travail est un peu organisé (section, titres, exemples) néanmoins le manque de détails et d'aspect mathématique rend la qualité du travail très médiocre.

2. Qualité du latex et des formules mathématiques : Il n'y a pas trop de formules mathématiques, donc on ne peut pas juger sa maîtrise du code latex.

3. Compréhension de l'idée générale : On comprend l'idée général du sujet mais juste parce que le sujet est très ordinaire et très simple, sinon l'aspect théorique ne suffit jamais pour comprendre ce genre de thème.

4. Explication et compréhension des formules mathématiques et des concepts fondamentaux: malheureusement on ne trouve pas de formules ou concepts mathématiques, ce qui rend l'évaluation à la base de ce critère impossible.

5. Difficulté et originalité du thème : le sujet abordé est très ordinaire, en termes de difficultés, c'est un sujet très facile.

6. Conclusion

le sujet est intéressant, Cependant, nous constatons que l'auteur n'a pas pu réaliser un document complet et riche de fondements mathématiques nécessaire à la compréhension des modèles de régression, ce qui nous laisse juger que le travail est malheureusement, médiocre.