

Finding the factors that directly affect time for passion

Guan Ming Chee

Machine Learning Final Project

Exploration of the Data

Introduction

- This project is trying to see the factors that are related to free time for passion.
- There are several features such as “flow”, “social network” and “live vision” being used to test the correlation with the time factor.

Rbf kernel

Cross validation confusion matrix are: [array([[0, 7],
[0, 21]], dtype=int64), array([[0, 5],
[0, 23]], dtype=int64), array([[1, 3],
[0, 24]], dtype=int64), array([[0, 4],
[0, 23]], dtype=int64), array([[0, 6],
[0, 21]], dtype=int64)]

Cross validation accuracies are: [0.75, 0.8214285714285714, 0.8928571428571429, 0.8518518518518519, 0.7777777777777778]

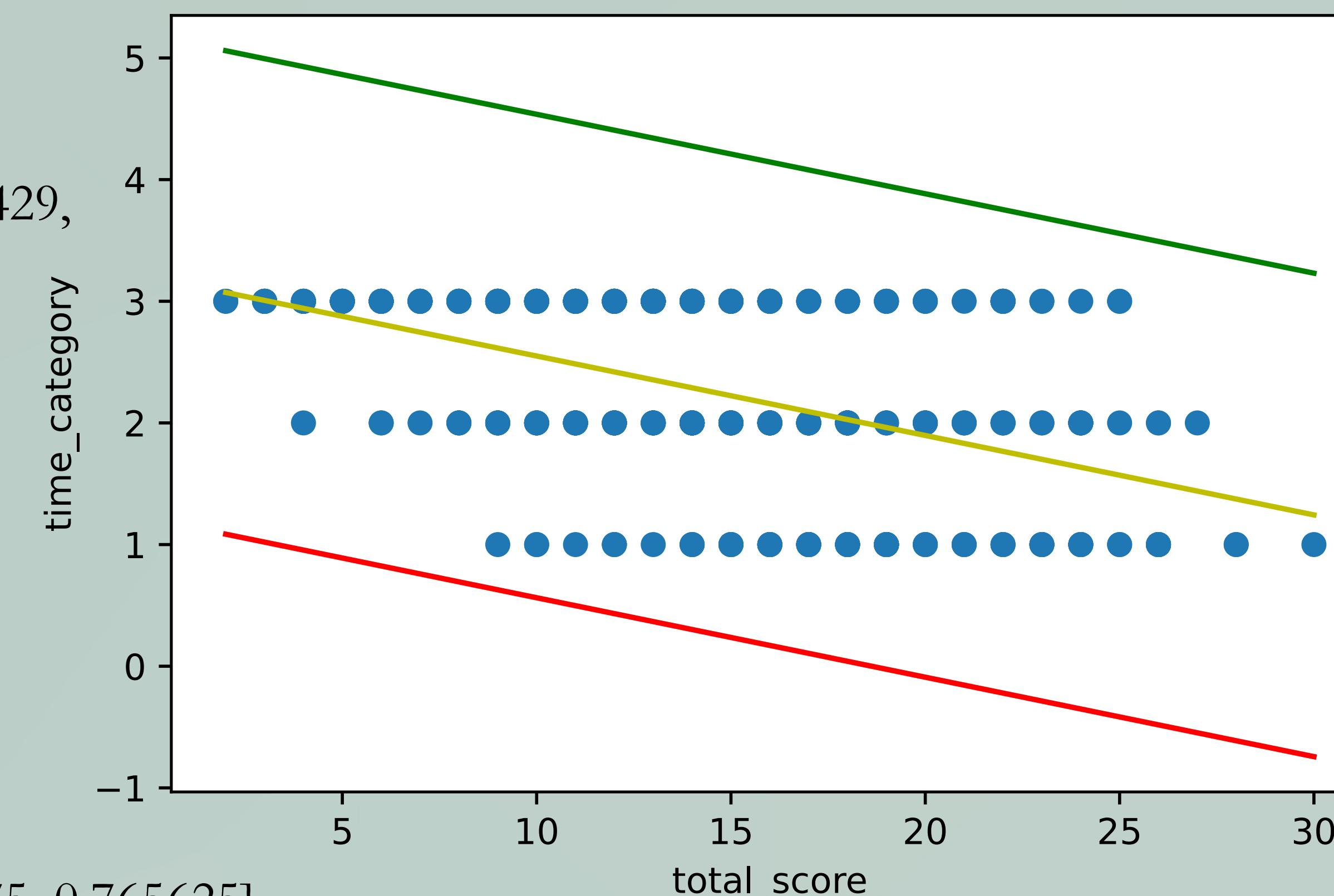
Cross validation f1 scores are: [0.6428571428571429, 0.7408963585434174, 0.8638655462184873, 0.7837037037037037, 0.6805555555555557]

Decision Tree

Cross validation confusion matrix are: [array([[7, 7],
[12, 39]], dtype=int64), array([[4, 11],
[12, 37]], dtype=int64), array([[6, 7],
[9, 42]], dtype=int64), array([[2, 8],
[13, 41]], dtype=int64), array([[3, 6],
[9, 46]], dtype=int64)]

Cross validation accuracies are: [0.7076923076923077, 0.640625, 0.75, 0.671875, 0.765625]

Cross validation f1 scores are: [0.7223031264268378, 0.6445689225141338, 0.7564285714285713, 0.6967233009708739, 0.7790804405874497]



What I found

- It can be understood that you need to get into the flow if you are doing something you are passionate about, and a well-planned life (life vision) helps to fulfilling your passion
- I used several kernels and RBF is the best among all. Decision Tree is suspected to be overfitting.
- The graph do not behave linearly or in polynomial shape as there are always exceptions/outliers among humans that makes prediction flawed.

Data

- FLOW: mental state, in which you are fully immersed in performing an activity.
- SOCIAL NETWORK: close family and friends ready to provide you with a long-term unconditional support.
- LIVE VISION: vision board, detailed in a personal journal or openly discussed with your spouse or close friends.
- TIME FOR PASSION: Daily hours spent doing what you are passionate
- SOURCE: <http://www.authentic-happiness.com/your-life-satisfaction-score>
- “FLOW”, “SOCIAL NETWORK” “LIVE VISION” AND “TIME FOR PASSION” are what I found to be correlated although the correlation are not strong.
- No data-cleaning is needed as the survey requires all fields to be filled.
- However, some datatypes needed to be changed because they do not belong to integer (int64) so comparison between numbers cannot be done.

Future Directions, Challenges and solutions

- More features can be added (CORE_CIRCLE, SUPPORTING_OTHERS, PERSONAL_AWARDS and ACHIEVEMENT)
- Newly added features tend to have weaker correlations which may drop the metrics score but we can try to limit the range (age) so that non-adults will not be affecting the prediction.
- Overall, human possess infinite possibilities so there will always be exceptions that is hard to tell and predict.

Poster Repository
GitHub

44-599-MachineLearning-S21/project-machine-learning-s21-GuanMingChee