

- 添付ファイルは、ある会社での社員の欠勤の記録です。
- 一番右のカラムが欠勤時間を示しています。今回は、10時間以上欠勤している人と、そうでない人を分類してください。
- 決定木とランダムフォレストを用いて、精度を比較してください。
 - データを8：2に分割し、学習データとテストデータとしてください。
 - テストデータの精度を示してください。
 - 属性IDは使わないようにしましょう。
 - クロスバリデーションの必要はありません。
 - どんなライブラリやツールを用いても構いませんが、pythonをおすすめします(scikit-learn中のライブラリ)。
- 締め切りは8/2(月)とします。

* English version is in next page.

- Attached file is a record of an employee's absences at a company.
- The rightmost column shows the hours of absence. In this case, classify those who are absent for more than 10 hours and those who are not.
- Use a decision tree and a random forest to compare the accuracy of the data.
 - Divide the data into 8:2, the training data and the test data.
 - Indicate the accuracy of the test data.
 - Do not use attribute ID.
 - There is no need for cross-validation.
 - You can use any library or tool you want. Here, we recommend python (scikit-learn).
- Deadline is 8/2(Mon).