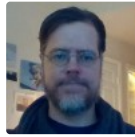


[Notebook](#) Search kaggle

Forks (2)

Fork Notebook

Sign



Brian Burns

Iris Exploration (PCA, k-Means and GMM clustering)

last run a year ago · Python notebook · 1097 views
using data from [Iris Species](#) · ...

3

voters

Notebook

Exploring the Iris dataset

This notebook explores the Iris dataset using visualization and **Principal Components Analysis (PCA)**, and compares **k-Means** and **Gaussian Mixture Model (GMM)** clustering using the **adjusted Rand score**.

Imports

```
In [1]: # import some libraries
import numpy as np
import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt

# seaborn is a layer on top of matplotlib which has additional visualizations -
# just importing it changes the look of the standard matplotlib plots.
# the current version also shows some warnings which we'll disable.
import seaborn as sns
sns.set(style="white", color_codes=True)
import warnings
warnings.filterwarnings("ignore")

# show plots inline
%matplotlib inline
```

Get Data

```
In [2]: # get data and show some records
data = pd.read_csv("../input/Iris.csv")
```

[Notebook](#)

[Code](#)

[Data \(1\)](#)

[Comments \(2\)](#)

[Log](#)

[Versions \(18\)](#)

[Forks \(2\)](#)

[Fork Notebook](#)