

به نام خدا

راهنمای استفاده از برنامه‌های کارگاه داده کاوی CSSS 2019

show\_columns(دیتاست)

نام تمامی ستون‌های دیتاست را به شما نشان می‌دهد.

drop\_na(دیتاست)

خانه‌های خالی دیتاست را پاک می‌کند.

summarize(لیستی به صورت گفته شده از نام ستون و یک عملیات , لیستی از نام ستون ها, دیتاست)

لیست : []

لیست قسمت سوم بدین صورت است : ['Title:count']

Title در این جا نام یکی از ستون های دیتاست و count عملیاتی است که قصد پیاده کردن آن را روی دیتاست دسته بندی شده بر اساس لیست قسمت دوم داریم.

filter\_records(شرط, دیتاست)

داده هایی از دیتاست که در شرط صدق میکند را جدا می کند.

شرط بدین گونه نوشته شود : 'column's name == /<= /... value'

view(شماره داده پایان , شماره داده شروع , دیتاست)

داده های بازه گفته شده را نمایش می دهد.

head(دیتاست)

۵ داده ی اول از دیتاست را نمایش می دهد.

tail(دیتاست)

۵ داده ی آخر از دیتاست را نمایش می دهد.

Max\_record(اسم ستون, دیتاست)

داده ای که بیشترین مقدار در ستون داده شده را دارد، نمایش می دهد.

min\_record(اسم ستون, دیتاست)

داده ای که کمترین مقدار در ستون داده شده را دارد، نمایش می دهد.

Sort(شرط, صعودی یا نزولی, لیستی از نام ستون ها, دیتاست)

داده را بر حسب اسم ستون های داده شده و شرط (میتوانید شرط را ورودی ندهید) به صورت صعودی یا نزولی مرتب می کند.

صعودی : True

نزولی : False

Unique\_records(اسم ستون, دیتاست)

داده های تکراری آن ستون را حذف و دیتاست جدید را چاپ می کند.

Agg(عملیات, نام ستون, دیتاست)

خروجی عملیات روی آن ستون را بر می گرداند برای مثال : min, max, mean

Apply(فرمول, نام ستون جدید, دیتاست)

در فرمول عملیاتی را روی دو ستون انجام داده، حاصل را در ستون جدید ذخیره می کند.

plot\_2d(رنگ, محور عمودی, محور افقی, دیتاست)

هر یک از قسمت های محور افقی، محور عمودی و رنگ نام یکی از ستون های دیتاست شما هستند.

نمودار دو بعدی رسم می کند.

Plot\_3d(رنگ, محور قائم, محور عمودی, محور افقی, دیتاست)

هر یک از قسمت های محور افقی، محور عمودی، محور قائم، رنگ نام یکی از ستون های دیتاست شما هستند.

نمودار سه بعدی رسم می کند.

Density(شرط, مقیاس, لیستی از ستون ها, دیتاست)

نمودار چگالی یک ستون که داده ی آنها در شرط صدق می کند را رسم می کند.

مقیاس توسط عددی به نام bin\_size مشخص می شود.

Bar\_chart(رنگ, محور عمودی, محور افقی, دیتاست)

هر یک از قسمت های محور افقی، محور عمودی، رنگ نام یکی از ستون های دیتاست شما هستند.

دیتاست را بر اساس ستون محور افقی دسته بندی کرده و تعداد ستون محور عمودی را برای هر دسته در یک نمودار میله ای نمایش می دهد.

Random\_grouping(نام ستون اطلاعات, نام ستون برای گروه بندی, دیتاست)

اطلاعات را به دو دسته دلخواه با اندازه های مشخص بر اساس ستون گروه بندی تقسیم کرده و نمودار جعبه ای دو گروه را رسم و میانگین هر گروه را نیز اعلام میکند.

Permutation\_test(نام ستون اطلاعات, نام ستون برای گروه بندی, دیتاست)

کار تابع قبلی را برای تعداد بار زیادی انجام داده و نمودار فراوانی تفاضل میانگین های دو گروه را رسم می کند.

Iris\_knn(نام ستون جهت پیش بینی, لیستی از ستون ها جهت پیش بینی, تعداد همسایه ها, دیتاست)

با استفاده از لیست داده شده ستون گفته شده را پیش بینی می کند. این تابع فقط برای دیتاست iris کار می کند.

(لیستی از ستون ها جهت پیش بینی ,نام ستون جهت پیش بینی ,تعداد همسایه ها ,دیتاست) KNN\_classifier

با استفاده از لیست داده شده ستون گفته شده را پیش بینی می کند. این تابع برای هر دیتاستی کار می کند.