

(1) این دیتا ست شامل اندازه گیری های پزشکی دقیقی است که برای پیش بینی شروع دیابت بر اساس چندین فاکتور سلامت جمع آوری شده است. این شامل 768 پرونده از بیماران زن است که هر کدام با 8 ویژگی مرتبط با سلامت مشخص می شوند. متغیر نتیجه نشان می دهد که آیا بیمار دیابت دارد (1) یا خیر (0). این مجموعه داده را می توان برای آموزش و آزمایش مدل های یادگیری ماشین برای کارهای طبقه بندی مرتبط با پیش بینی دیابت استفاده کرد و برجسب این دیتا ست ستون Outcome است.

(2) تمام ویژگی های ذکر شده در دیتا ست عددی (Numerical) هستند به جز ستون آخر که به صورت دودویی یا (Binary) است.

(3) بله، کد آن در فایل موجود است.

(4) کد در فایل موجود است.

(5) نمودار هیستوگرام این دیتا ست.

