

# K nearest neighbors

K نزدیکترین همسایه (knn) یک مدل یادگیری ماشینی نظارت شده است که یک نقطه داده را می گیرد، به نزدیک ترین نقاط داده برچسب گذاری شده 'k' خود نگاه می کند و با اکثریت آرا برچسب را اختصاص می دهد.

در اینجا می بینیم که تغییر k می تواند بر خروجی مدل تأثیر بگذارد. در k، knn یک فرایارامتر است. فرایارامتر در یادگیری ماشینی پارامتری است که مقدار آن قبل از شروع فرآیند یادگیری تنظیم می شود. در ادامه نحوه تنظیم هایپر پارامتر را یاد خواهیم گرفت.

به عنوان مثال، در شکل زیر دو کلاس وجود دارد: مربع های آبی و مثلث های قرمز. بر اساس الگوریتم 3nn، یعنی زمانی که k 3 است، چه برچسبی را باید به نقطه سبز با برچسب ناشناخته اختصاص دهیم؟ از 3 نزدیکترین نقطه داده از نقطه سبز (دایره خط یکدست)، دو مثلث قرمز و یکی مربع آبی است، بنابراین نقطه سبز به عنوان یک مثلث قرمز پیش بینی می شود. اگر k 5 باشد (دایره خط چین)، سپس به عنوان مربع آبی طبقه بندی می شود (3 مربع آبی در مقابل 2 مثلث قرمز، مربع های آبی اکثریت هستند).

