DBSCAN

Rafia Amin Raka

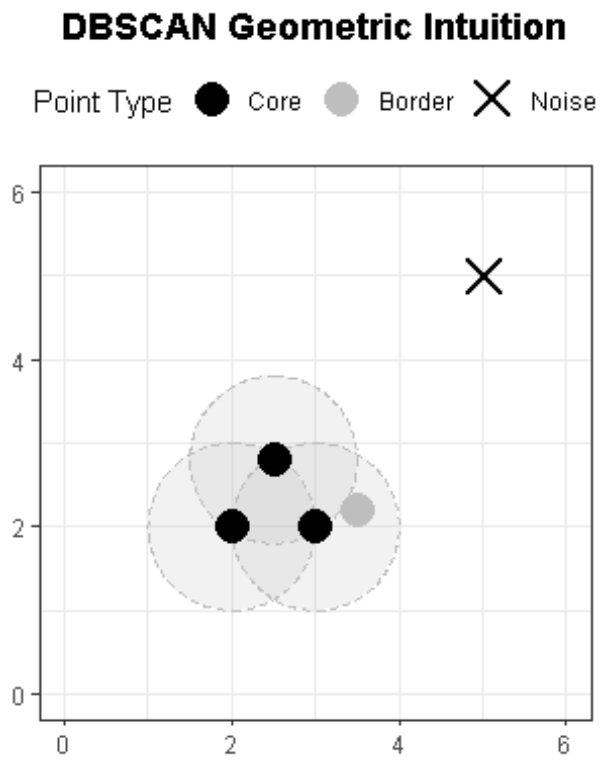
2025-08-14

```r
# install.packages("ggplot2")
# install.packages("ggforce")

library(ggplot2)

library(ggforce)

points_df <- data.frame(
  x = c(2, 3, 2.5, 3.5, 5),
  y = c(2, 2, 2.8, 2.2, 5),
  type = factor(c("Core", "Core", "Core", "Border", "Noise"),
                levels = c("Core", "Border", "Noise"))
)
points_df
```

```
##     x   y   type
## 1 2.0 2.0   Core
## 2 3.0 2.0   Core
## 3 2.5 2.8   Core
## 4 3.5 2.2 Border
## 5 5.0 5.0  Noise
```

```r
epsilon <- 1.0
core_points_circles <- data.frame(
  x0 = c(2, 3, 2.5),
  y0 = c(2, 2, 2.8),
  r = epsilon
)


ggplot() +
  # Draw the epsilon circles for core points
  geom_circle(
    data = core_points_circles,
    aes(x0 = x0, y0 = y0, r = r),
    color = "gray",
    fill = "gray",
    alpha = 0.2,
    linetype = "dashed"
```

```
  ) +
  geom_point(
    data = points_df,
    aes(x = x, y = y, shape = type, color = type),
    size = 5,
    stroke = 1.5 # Makes shapes like 'x' thicker
  ) +
  scale_color_manual(values = c(Core = "black", Border = "gray", Noise =
"black")) +
  scale_shape_manual(values = c(Core = 16, Border = 16, Noise = 4)) + # 16 is
a solid circle, 4 is an 'x'
  coord_fixed(xlim = c(0, 6), ylim = c(0, 6)) +
  labs(
    title = "DBSCAN Geometric Intuition",
    x = NULL, y = NULL, color = "Point Type", shape = "Point Type"
  ) +
  theme_bw() +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold"),
    legend.position = "top"
  )
```



Interpretation

The DBSCAN Geometric Intuition clearly distinguishes between core points, border points, and noise, using spatial relationships and neighborhood radius (ε) to illustrate the

concept. The core points, surrounded by overlapping circles, show how dense regions are formed, while the border point and isolated noise point emphasize DBSCAN's ability to handle outliers and irregular cluster shapes. This intuitive representation makes it easier to grasp DBSCAN's advantages over other clustering methods, especially its flexibility in discovering clusters without needing to predefine their number.