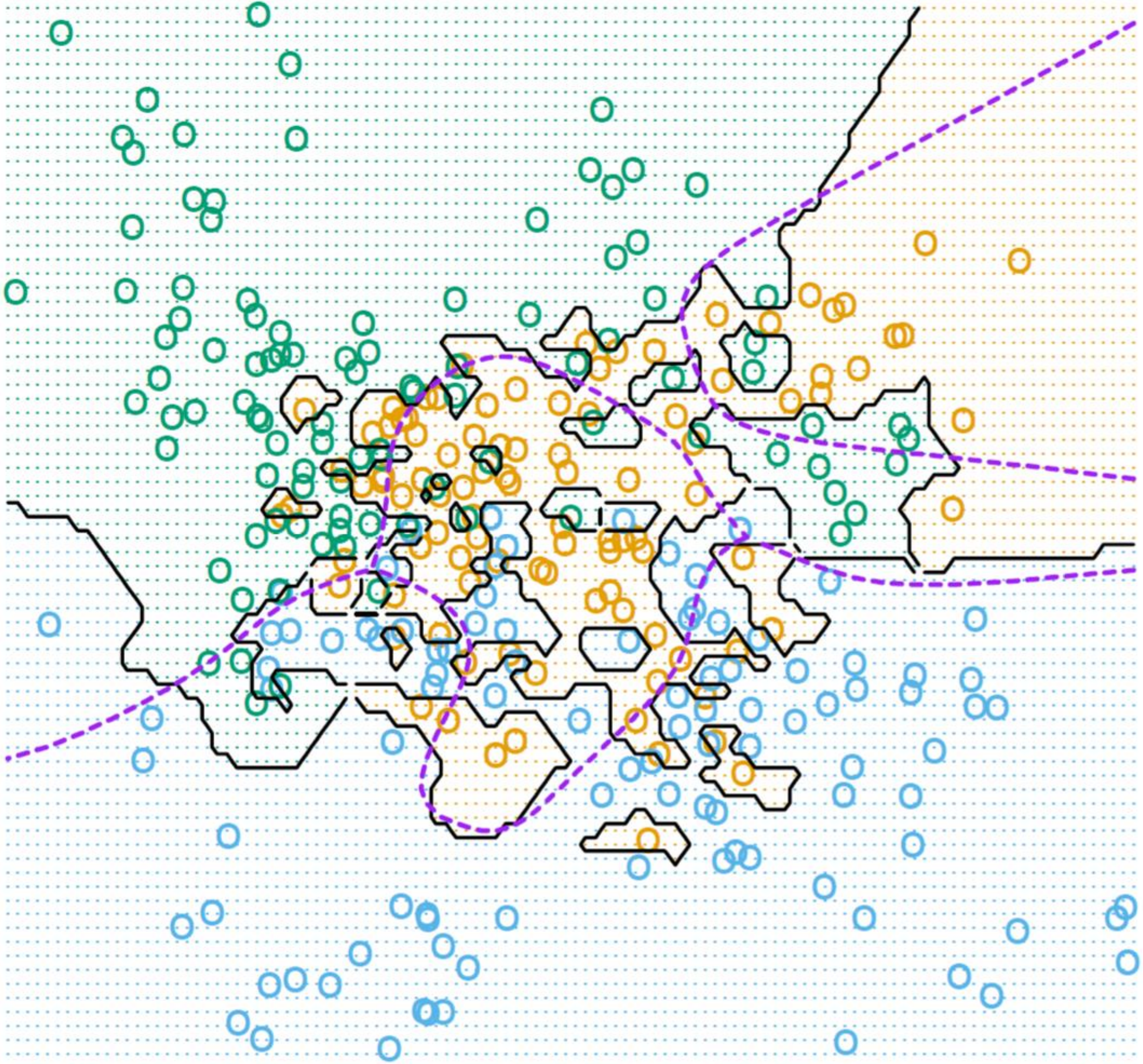
# NEAREST NEIGHBOUR CLASSIFIER

Lecture-4,5



# CLASSIFICATION BOUNDARY WITH NNC

- 2-d Euclidean example
- Dirichlet Tesselation or Voronoi Decomposition
- Bayes optimal in purple

Copied from: Tibshirani book

# ASYMPTOTIC BOUND

- **Theorem:** Let $R^{NN}$,and $R^*$denote the limiting value of the expected misclassification error with NNC (i.e., $R^{NN} = \lim_{m\to\infty} E[P[Y \neq g_m^{NN}(X) \mid X]]$) and with Bayes optimal respectively, then, under mild conditions, we have

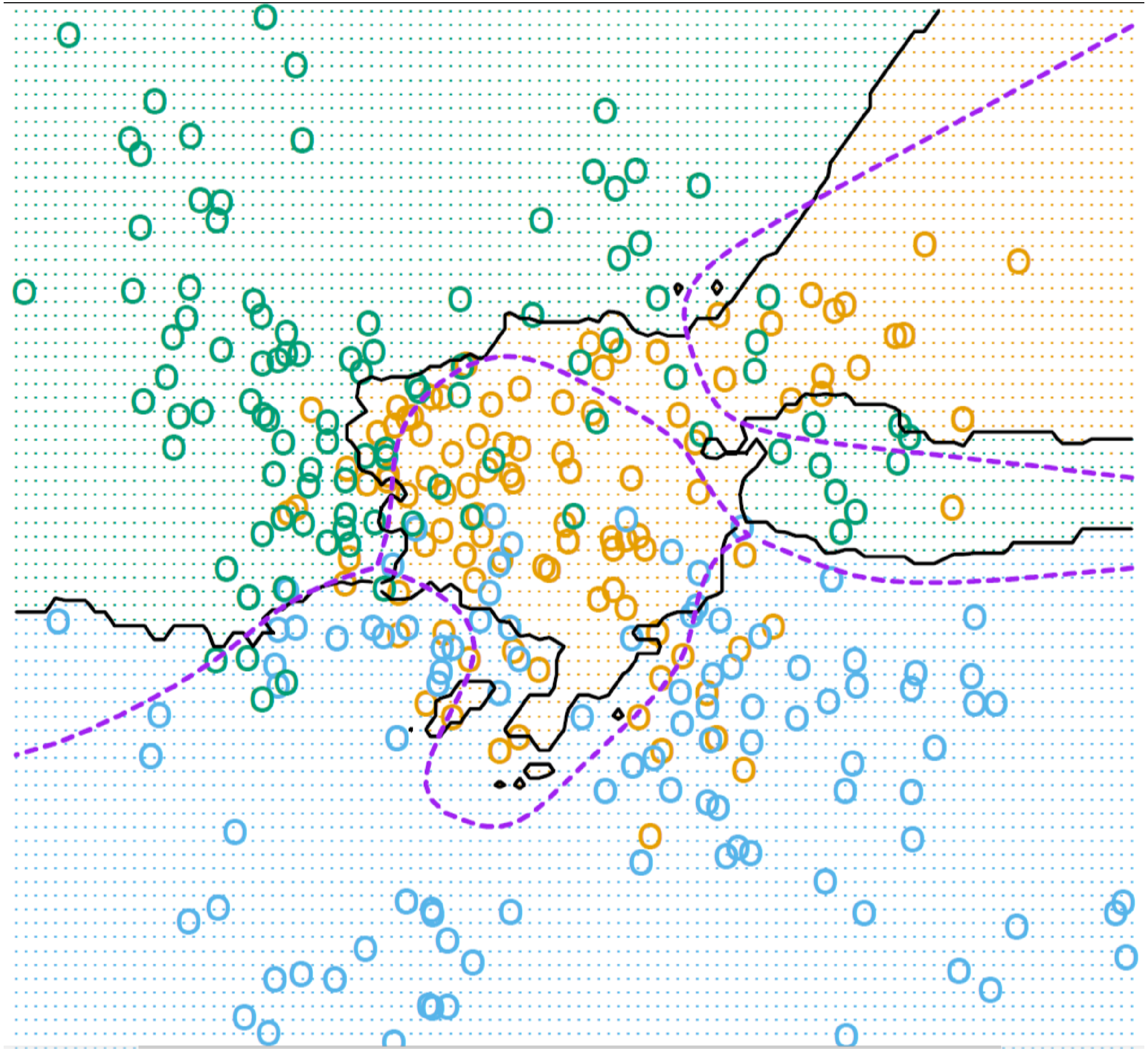
$$0 \leq R^* \leq R^{NN} \leq 2R^*$$

# ASYMPTOTIC BOUND

- **Theorem:** Let $R^{NN}$,and $R^*$denote the limiting value of the expected misclassification error with NNC (i.e., $R^{NN} = \lim_{m\to\infty} E[P[Y \neq g_m^{NN}(X) \mid X]]$) and with Bayes optimal respectively, then, under mild conditions, we have

$$0 \leq R^* \leq R^{NN} \leq 2R^*$$

- Interesting extreme cases:
  - $R^* = 0$, then $R^{NN} = 0$
  - $R^* = 0.5$, then $R^{NN} = 0.5$



# CLASSIFICATION BOUNDARY WITH KNNC

- Smoother boundary
- Closer to Bayes optimal

Copied from: Tibshirani book

# ASYMPTOTIC BOUND

- **Theorem:** Let $R^{kNN}$,and $R^*$denote the limiting value of expected misclassification rate with k-NNC (as $m \to \infty$) and that with Bayes optimal respectively, then, under mild conditions, we have

$$0 \leq R^* \leq R^{(2k+1)NN} \leq R^{(2k-1)NN} \leq 2R^*$$

- However, with finite $m$, it may happen that $R_m^{(2k+1)NN} \geq R_m^{(2k-1)NN}$

# BAYES CONSISTENCY [STONE77]

- **Theorem:** For large enough $m, k$, such that $\frac{k}{m}$ is small enough, we have with probability atleast $1 - \delta$:

$$0 \leq R_m^{kNN} - R^* \leq \sqrt{\frac{72\gamma^2 \log 2/\delta}{m}}$$

- In particular, in the limit $\mathrm{m} \rightarrow \infty, k \rightarrow \infty, \frac{k}{m} \rightarrow 0$, we have that $\boldsymbol{R^{NN} = R^*}$ (irrespective of the value of $R^*$)

# BAYES CONSISTENCY [STONE77]

- **Theorem:** For large enough $m, k$, such that $\frac{k}{m}$ is small enough, we have with probability atleast $1 - \delta$:

$$0 \leq R_m^{kNN} - R^* \leq \sqrt{\frac{72\gamma^2 \log 2/\delta}{m}}$$

$$\gamma \leq \left(1 + \frac{2}{\sqrt{2 - \sqrt{3}}}\right)^n - 1$$

Curse of dimensionalty (here artefact of analysis)