RESEARCH PAPER #1 REFERENCE: Classifying post-traumatic stress disorder using the magnetoencephalographic connectome and machine learning

RESEARCH FOCUS:

- → Developing an objective method to diagnose/distinguish combat-related PTSD.
- → Implementing a machine learning framework including support vector machines (SVM) for classification.
- → Data from MEG scans from individuals with combat-related PTSD.

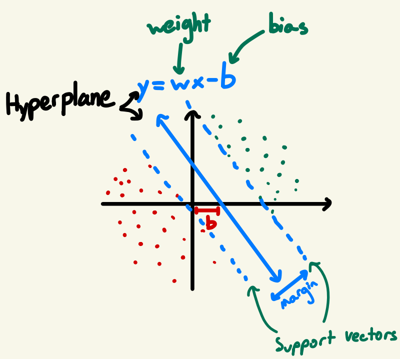
SUPPORT VECTOR MACHINES

- → A SVM classifier works well with higher dimensional data or with smaller data sets.
- → Works well with a CLEAR MARGIN OF SEPERATION

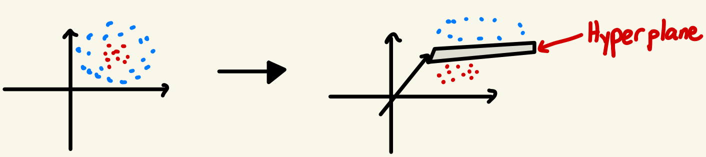
Makes me wonder how do we know if our models have a clear seperation of data?

NOTES ON SVM:

Hyperplane
↑
Line/Plane seperating data into 2 classes



→ Use a mathematical function known as **Kernal** to increase the dimensional space to better seperate data



→ Calculate the label. The sign of the output determines its classification.

• weight = (m, c) ← intercept of hyperplane

↑ slope of hyperplane

Label = $w^T x_i + b = 0$ ← Equation of Hyperplane

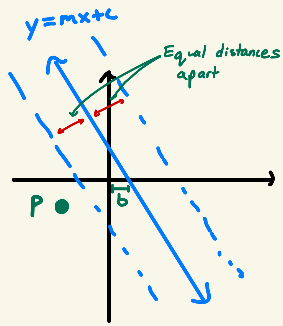
$w^T x + b = -1$

$w^T x + b = +1$

Support vectors

$w^T x = \begin{bmatrix} m \\ c \end{bmatrix} \cdot \begin{bmatrix} x & y \end{bmatrix}$

↑ Dot Product



Group

$$y \begin{cases} +1 & \text{if } \; y - mx - b \geq 0 \\ -1 & \text{if } \; y - mx - b < 0 \end{cases}$$

PREDICTION!

→ The best hyperplane has the greatest margin.

$$\because d = \frac{|c_2 - c_1|}{\sqrt{A^2 + B^2}} = \frac{|c_2 - c_1|}{||w||}$$

$$\therefore \max \frac{|c_2 - c_1|}{||w||} \text{ such that } w^T x_i + b \; \begin{matrix} \geq 1 \text{ if } Y_i = +1 \\ \leq -1 \text{ if } Y_i = -1 \end{matrix}$$

To maximize margin, we minimize $||w||$; however, we can risk incorrectly classifying data.

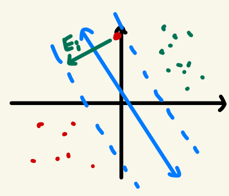
∴ When creating hyperplanes, we take into account incorrectly classified data.

$$\therefore \min \cdot |c_2 - c_1| \cdot ||w||^2 + C \sum_i^N \varepsilon_i$$

Distance measured from corresponding support vectors

↑ Known as "Slack variable"

Tuning parameter that controls how much weight to put on incorrectly classified information



→ Gradient Descent is an optimization algorithm used to minimize the cost function to find optimum parameters

↑

We use it to update the parameters of the model

$$w_2 = w_1 - L \cdot \frac{dJ}{dw}$$

$$b_2 = b_1 - L \cdot \frac{dJ}{db}$$

w = weight

b = bias

L = Learning Rate

$\frac{dJ}{dw}$ = Derivative of cost function with respect to w

$\frac{dJ}{db}$ = Derivative of cost function with respect to b

IF: $y_i \cdot (w.x + b) \geq 1$

$$\frac{dJ}{dw} = 2\lambda w$$

$$\frac{dJ}{db} = 0$$

ELSE: $y_i \cdot (w.x + b) < 1$

$$\frac{dJ}{dw} = 2\lambda w - y_i \cdot x_i$$

$$\frac{dJ}{db} = y_i$$

Regularization parameter

(controls incorrectly classified information)

RESULTS:

- There is distinct patterns of neural synchrony that were found across the five frequency bands (delta, theta, alpha, beta & gamma)

- The CV-SVM-rRf-FS approach minimized the number of features required for classification, reducing overfitting

- The retained edges (features) included:
  - Frontal Cortex
  - Hippocampus
  - Amygdala
  - Thalamus

Brain regions associated with PTSD

Suggests that these regions are ALTERED in PTSD