CS/ECE/ME 532 Period 24

Last day of class

- Finish kernel methods

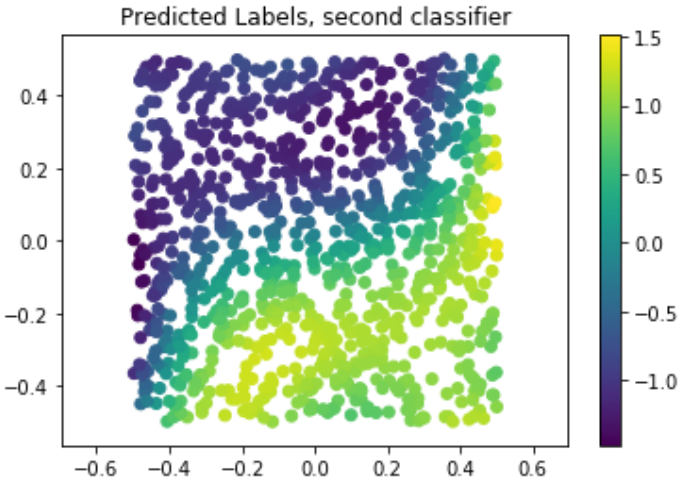
Remaining …

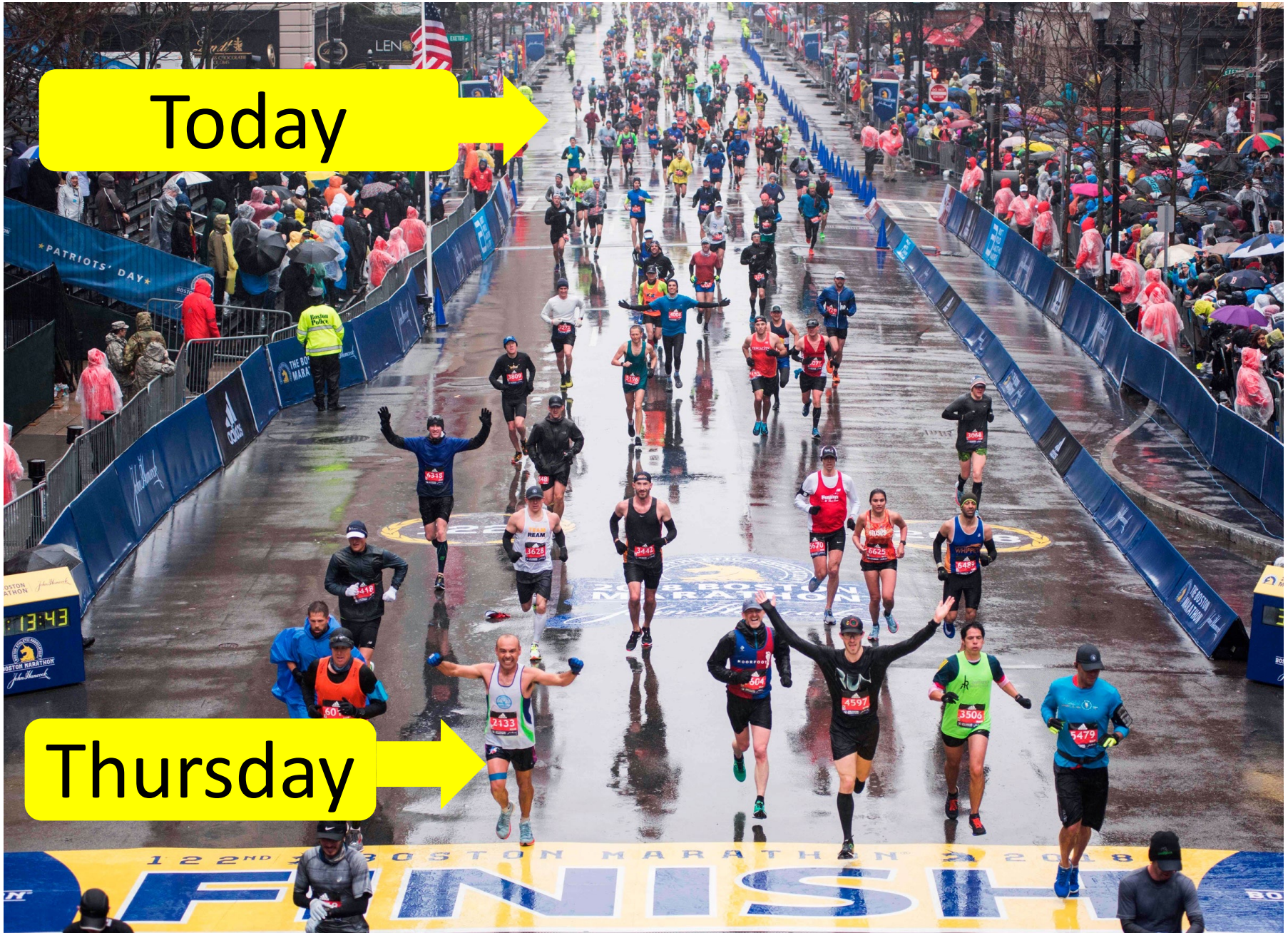
- Homework 9 due Wednesday night
- Unit 6 Quiz Thursday in class

- Unit 6 Integrative Summary
- Course Integrative Summary

Applications of Kernel methods

- *Classification using Kernel Methods*




Today
Thursday
PATRIOTS' DAY
FINISH
BOSTON MARATHON

Kernels for classification

[Kimeldorf, Wahba 1970]

https://www.uwalumni.com/news/assigned-reading-grace-wahba/

Classification, after feature map:

$$\widehat{y} = \text{sign}(\boldsymbol{\phi}(\boldsymbol{x})^T \boldsymbol{w}) \qquad (1)$$

$\boldsymbol{w}$ depends on $\boldsymbol{x}_1, y_1, \boldsymbol{x}_2, \boldsymbol{y}_2 \ldots$

Representer Theorem: (1) and (2) are the same, when . . .

$$\boldsymbol{w}^* = \arg\min_{\boldsymbol{w}} ||\boldsymbol{\Phi}\boldsymbol{w} - \boldsymbol{y}|| + \lambda ||\boldsymbol{w}||^2$$

$$\boldsymbol{w}^* = (\boldsymbol{\Phi}^T \boldsymbol{\Phi} + \lambda \boldsymbol{I})^{-1} \boldsymbol{\Phi}^T \boldsymbol{y}$$

and

$$\boldsymbol{\alpha} = (K + \lambda \boldsymbol{I})^{-1} \boldsymbol{y}$$

where K has ℓ, m entry $K(\boldsymbol{x}_\ell, \boldsymbol{x}_m)$

Kernel methods – re-write above as:

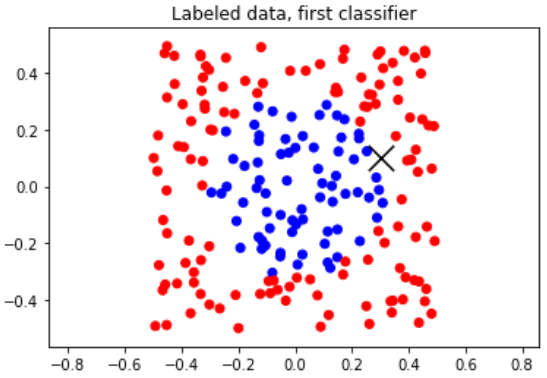
$$\widehat{y} = \text{sign}\left(\sum_i \alpha_i K(\boldsymbol{x}, \boldsymbol{x}_i)\right) \qquad (2)$$

weighted sum of similarities between feature vector and each training point

Example of kernel classification

$$\boldsymbol{x} = \begin{bmatrix} 0.3 \\ 0.1 \end{bmatrix}$$

How do we predict class of $\boldsymbol{x}$?



$$K(\boldsymbol{x}, \boldsymbol{x}_i) = \exp\left(-||\boldsymbol{x} - \boldsymbol{x}_i||^2\right)$$

$$\widehat{y} = \text{sign}\left(\sum_i \alpha_i \exp\left(-\left|\left|\begin{bmatrix} 0.3 \\ 0.1 \end{bmatrix} - \boldsymbol{x}_i\right|\right|^2\right)\right)$$