Naive Bayes (continued)

(find and work with a partner)

1. Continuing the example from the previous handout, say we have a new data point $\boldsymbol{x}_{\text{test}} = [\text{neg}, \text{pos}]$. Our goal is to predict the class label based on the Naive Bayes posterior probability. In practice, we will compute this probability for each class k, based on our estimates (θ_k and $\theta_{k,j,v}$ terms). Then we will assign this data point the class label with maximum probability:

$$\hat{y} = \underset{k \in \{1,2,\cdots,K\}}{\arg\max}\ p(y = k|\boldsymbol{x}) = \underset{k \in \{1,2,\cdots,K\}}{\arg\max}\ p(y = k) \prod_{j=1}^{p} p(x_j|y = k).$$

   For this $\boldsymbol{x}_{\text{test}}$, compute $p(y = 1|\boldsymbol{x}_{\text{test}})$ and $p(y = 2|\boldsymbol{x}_{\text{test}})$ and then assign a prediction label $\hat{y}$.

2. For the tennis example below, fill in the $\theta_{k,j,v}$ terms. Discuss how these values could be implemented using dictionaries.

Day	Outlook	Temperature	Humidity	Wind	PlayTennis (y)
$\boldsymbol{x}_1$	Sunny	Hot	High	Weak	No
$\boldsymbol{x}_2$	Sunny	Hot	High	Strong	No
$\boldsymbol{x}_3$	Overcast	Hot	High	Weak	Yes
$\boldsymbol{x}_4$	Rain	Mild	High	Weak	Yes
$\boldsymbol{x}_5$	Rain	Cool	Normal	Weak	Yes
$\boldsymbol{x}_6$	Rain	Cool	Normal	Strong	No
$\boldsymbol{x}_7$	Overcast	Cool	Normal	Strong	Yes
$\boldsymbol{x}_8$	Sunny	Mild	High	Weak	No
$\boldsymbol{x}_9$	Sunny	Cool	Normal	Weak	Yes
$\boldsymbol{x}_{10}$	Rain	Mild	Normal	Weak	Yes
$\boldsymbol{x}_{11}$	Sunny	Mild	Normal	Strong	Yes
$\boldsymbol{x}_{12}$	Overcast	Mild	High	Strong	Yes
$\boldsymbol{x}_{13}$	Overcast	Hot	Normal	Weak	Yes
$\boldsymbol{x}_{14}$	Rain	Mild	High	Strong	No

y=No (0)

outlook	**Sunny:**	**Overcast:**	**Rain:**
temperature	**Cool:**	**Mild:**	**Hot:**
humidity	**Normal:**	**High:**	
wind	**Weak:**	**Strong:**	

y=Yes (1)

outlook	**Sunny:**	**Overcast:**	**Rain:**
temperature	**Cool:**	**Mild:**	**Hot:**
humidity	**Normal:**	**High:**	
wind	**Weak:**	**Strong:**	