SET A

Q1

Let $X = \{x_1, x_2\}$ be the age groups and $Y = \{y_1, y_2\}$ indicate if a person is a subscribe or not. Let x_1 be Bob's age group and y_1 indicate that a person is a subscriber. It is given that the priors are equal, hence $P(y_1) = 0.5$ and $P(y_2) = 0.5$.

$$
\begin{aligned}
P(y_1 \mid x_1) &= \frac{P(y_1) \times P(x_1 \mid y_1)}{P(x_1)} \\
0.8 &= \frac{P(y_1) \times P(x_1 \mid y_1)}{P(x_1 \mid y_1)P(y_1) + P(x_1 \mid y_2)P(y_2)} \\
0.8 &= \frac{P(x_1 \mid y_1)}{P(x_1 \mid y_1) + P(x_1 \mid y_2)} \text{ , since, } P(y_1) = P(y_2) = 0.5
\end{aligned}
$$

Hence, let $P(x_1 \mid y_1) = 0.8k$ and $P(x_1 \mid y_2) = 0.2k$, where k is some fixed constant.

$$
\begin{aligned}
P(y_1 \mid \{x_1, x_1\}) &= \frac{P(y_1) \times P(\{x_1, x_1\} \mid y_1)}{P(\{x_1, x_1\})} \\
&= \frac{P(\{x_1, x_1\} \mid y_1)}{P(\{x_1, x_1\} \mid y_1) + P(\{x_1, x_1\} \mid y_2)} \\
&= \frac{P(x_1 \mid y_1)^2}{P(x_1 \mid y_1)^2 + P(x_1 \mid y_2)^2} \\
&= \frac{(0.8k)^2}{(0.8k)^2 + (0.2k)^2} \\
&= 0.94
\end{aligned}
$$

Q2

[slide 27 in 6-3 classification2.pdf]

Q3

Bagging. For bagging, the models can be trained in parallel, not in boosting (in boosting the models have to be trained sequentially). Hence, 10 models can be trained in parallel on the first and 10 on second day, with one extra day to spare.

Q4

[slide 29 in 7 apaug ensembles-short.pdf]
Properties that make XGBoost a popular ensemble model are -

1. Parallel Processing utilizing multiple CPU cores
2. Can be used for both classification as well as regression tasks
3. Supports a distributed implementation on Hadoop

The challenge in training an XGBoost model is hyperparameter tuning of a large number of parameters.