机器学习课程　第八次作业

黄昊　　20204205

选择题目： 8.1　8.4　8.6　8.7　8.8

8.1　证明：

$$
\begin{aligned}
P(H(x) \neq f(x)) &= P(X \leq \lfloor T/2 \rfloor) \\
&\leq P(X \leq T/2) \\
&= P(X - (1-\epsilon)T \leq \frac{T}{2} - (1-\epsilon)T) \\
&= P(X - (1-\epsilon)T \leq -\frac{T}{2}(1-2\epsilon)) \\
&= P(\sum_{i=1}^{T} x_i - \sum_{i=1}^{T} \mathbb{E}(x_i) \leq -\frac{T}{2}(1-2\epsilon)) \\
&= P(\frac{1}{T}\sum_{i=1}^{T} x_i - \frac{1}{T}\sum_{i=1}^{T} \mathbb{E}(x_i) \leq -\frac{1}{2}(1-2\epsilon))
\end{aligned}
$$

由Hoeffding不等式可知：

$$
P(\frac{1}{m}\sum_{i=1}^{m} x_i - \frac{1}{m}\sum_{i=1}^{m} \mathbb{E}(x_i) \leq -\delta) \leq e^{-2m\delta^2}
$$

令$\delta = \frac{1-2\epsilon}{2}, m = T$得：

$$
\begin{aligned}
P(H(x) \neq f(x)) &= \sum_{k=0}^{\lfloor T/2 \rfloor} \binom{T}{k} (1-\epsilon)^k \epsilon^{T-k} \\
&\leq e^{-\frac{1}{2}T(1-2\epsilon)^2}
\end{aligned}
$$

从而得证。

8.4　两种算法都属于集成学习算法。但Adaboost是通过增加上一轮训练错误样本的权重达到关注预测错误样本的目的，而GradientBoosting 是用负梯度来作为上一轮中基学习器犯错的衡量指标，从而在下一轮中通过拟合上一轮中的负梯度来达到纠正上一轮中所犯错误的目的。两者的优化对象有所区别。

8.6　根据偏差-方差分解，Bagging关注于降低方差，而朴素贝叶斯是在所有样本中求最大后验概率，某种程度上已经是考虑样本后的最优解，此时再降低方差是没有意义的。

8.7　因为随机森林除了在样本的选择是随机选择之外，在属性上的选择也是随机选取一部分，增加了基学习器的多样性。

8.8　Multiboosting 可以有效降低误差和方差，但训练成本会显著上升。 Iterative Bagging会通过bagging降低误差，但是方差会上升。