教务处填写：

___年___月___日
考 试 用

湖南大学课程考试试卷

课程名称： 机器智能 ；**课程编码：** CS06152 ；

试卷编号： A ；**考试形式：** 闭卷 ；**考试时间：** **120** 分钟。

题 号	一	二	三	四	五	六	七	八	九	十	总分
应得分	20	25	20	15	20						100
实得分											
评卷人											

湖南大学课程考试试卷

专业班级：

学号：

姓名：

湖南大学教务处

装订线（题目不得超过此线）

（请在答题纸内作答！）

一、（简答题）（18 分）

(a). 以图像分类为例，请简述卷积神经网络最主要的三个操作以及每个操作的主要作用？（9 分）

卷积层：采用卷积操作来提取图像的卷积特征，其主要的作用是图像语义特征提取；（3 分）

池化层：采用池化操作来对卷积后的特征进行降维，其主要作用是在保留图像信息的同时降低特征维度；（3 分）

全连接层：根据提取的特征进行预测输出值，其主要的作用是预测图像分类结果。（3 分）

(b). 比较有监督学习，无监督学习和半监督之间的差异，并简述三种学习方式的优点；（9 分）

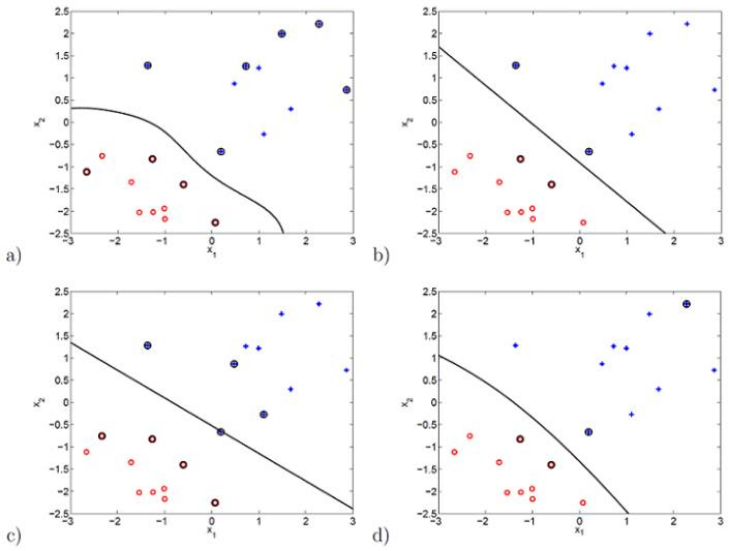
有监督学习采用带标签的数据进行学习；（2 分）其主要的优点是性能较好；（1 分）

无监督学习采用不带标间的数据进行学习；（2 分）其主要的优点是节约了数据标注的时间（1 分）

半监督学习融合了两者的特点，既使用带标签的数据又使用不带标签的数据进行学习；（2 分）其主要的优点是既保证了进行又节约了人工标注的代价（1 分）

二、（支持向量机）（20 分）

下图为采用不同核函数或不同的松弛因子得到的 SVM 决策边界。但粗心的实验者忘记记录每个图形对应的模型和参数了。请你帮忙给下面每个模型标出正确的图形并解释原因。



(a) 、 $\min\left(\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{i=1}^{N}\xi_i\right)$, s.t. $\xi_i\geq 0,\ y_i\left(\mathbf{w}^T\mathbf{x}+w_0\right)\geq 1-\xi_i,\ i=1,....,N,$ 其中$C=0.1$。

图 C; （3 分）线性分类面，C 较小，正则较大，||w||较小，Margin 较大, 支持向量较多（2 分）

(b)、 $\min\left(\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{i=1}^{N}\xi_i\right)$, s.t. $\xi_i\geq 0,\ y_i\left(\mathbf{w}^T\mathbf{x}+w_0\right)\geq 1-\xi_i,\ i=1,....,N,$ 其中$C=1$。

图 b; （3 分）线性分类面，C 较大，正则较小，||w||较大，Margin 较小支持向量的数目少（2 分）

(c)、 $\max\left(\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j k\left(\mathbf{x}_i,\mathbf{x}_j\right)\right)$ s.t. $0\leq\alpha_i<C,\ i=1,....,N,$ $\sum_{i=1}^{N}\alpha_i y_i=0$

其中$k\left(\mathbf{x},\mathbf{x}'\right)=\mathbf{x}^T\mathbf{x}'+\left(\mathbf{x}^T\mathbf{x}'\right)^2$。

图 d; （3 分）二次多项式核函数，决策边界为二次曲线 （2 分）

（d）、 $\max\left(\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j k\left(\mathbf{x}_i,\mathbf{x}_j\right)\right)$ s.t. $0\leq\alpha_i<C,\ i=1,....,N,$ $\sum_{i=1}^{N}\alpha_i y_i=0$

其中$k\left(\mathbf{x},\mathbf{x}'\right)=\exp\left(-\frac{1}{2}\|\mathbf{x}-\mathbf{x}'\|^2\right)$。

图 a; （3 分）RBF 核函数，决策边界为曲线，σ =1 较大，曲线更平滑（2 分）

三、（搜索问题）（25 分）

考虑如下的游戏：有三个黑色的地砖 B, 三个白色的地砖 W 以及一个空的位置⌀，初始的状态为<B|B|B|W|W|W|⌀>, 游戏规则如下：

1. B 或者 W 可以移到邻居位置上去当且仅当邻居位置为空；
2. B 或者 W 可以跳过 1-2 个地砖到达一个空的位置；

3. 移动一个地砖的代价为 1 加上移动的距离（比如，移动到邻居位置的代价为 1，隔一个位置移动的代价为 2，最大的代价为 3）；

4. 游戏的目标是 W 都位于 B 的左边（不关心空的位置）；

请回答下列问题：

(a) 针对该问题设计一种启发式的函数；（10 分）

比如 B 位于 W 左边的数量等等，答案不唯一

(b) 根据(a)中的启发式函数使用 A*树算法进行问题求解，要求描述具体步骤 （15 分）。

答案不唯一

四、（贝叶斯网络）（20 分）

考虑如下的贝叶斯网络“Asia”用来诊断呼吸疾病，回答如下问题：

（a）计算联合概率 P(a, s, t, ~l, b, e, x, d)；（6 分）

$P(a, s, t, \sim l, b, e, x, d)=0.01*0.5*0.05*0.9*0.6*1*0.98*0.9=1.1907*10^{-4}$

（b）计算(l | s, a, x) .（8 分）

P(l|s,a,x)=a*P(l|s)sum_{T,E}[P(T|a)*P(E|T,l)*P(x|E)]

=0.1a*[0.05*1*0.98+0.95*1*0.98]

=0.1a*0.98=0.098a

P(~l|s,a,x)=a*P(~l|s)sum_{T,E}[P(T|a)*P(E|T,~l)*P(x|E)]

=0.9a*[0.05*1*0.98+0.95*1*0.05]

=0.08685a

归一化后得到 P(l|s,a,x)=0.53

（c） 判断如下变量之间的独立性 .（6 分）

给定 a，l 和 t 之间是否独立；

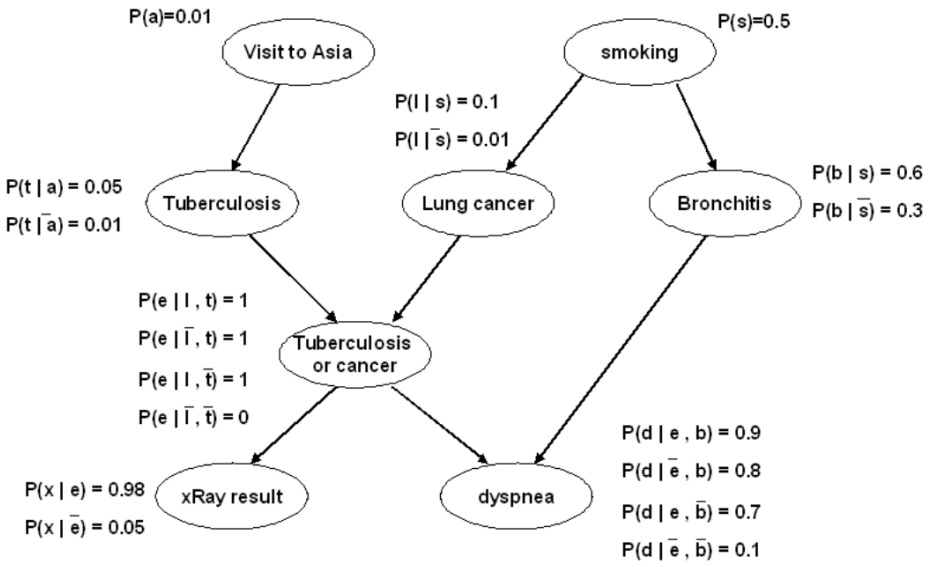
独立 （2 分）

给定 a 和 e,l 和 t 之间是否独立；

不独立 （2 分）

给定 s 和 e， l 和 b 之间是否独立；

独立 （2 分）



五、（决策树）（17 分）

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

(a) . 根据上表构造一颗决策树，其中 playtennis 是标签，要求计算属性的信息增益值（12 分）；

The initial entropy of the training sample:

$$E(S) = -(\frac{5}{14}log_2\frac{5}{14} + \frac{9}{14}log_2\frac{9}{14}) = 0.9403$$

We will choose the variable to split on such that the corresponding information gain is maximal.

$$InfoGain = E(S) - \sum_{vals}\frac{\mid S_v \mid}{\mid S \mid}E(S_v)$$

$$InfoGain(S,T) = 0.9403 - \frac{4}{14} - \frac{6}{14}(-(\frac{2}{6}log_2\frac{2}{6} + \frac{4}{6}log_2\frac{4}{6})) - \frac{4}{14}(-(\frac{1}{4}log_2\frac{1}{4} + \frac{3}{4}log_2\frac{3}{4})) = 0.0292$$

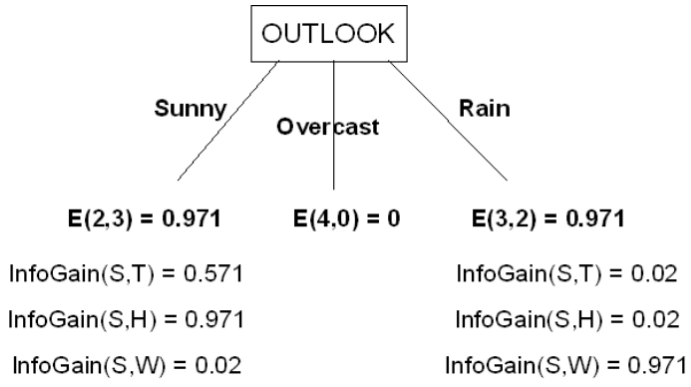
$$InfoGain(S,H) = 0.9403 - \frac{7}{14}(-(\frac{4}{7}log_2\frac{4}{7} + \frac{3}{7}log_2\frac{3}{7})) - \frac{7}{14}(-(\frac{6}{7}log_2\frac{6}{7} + \frac{1}{7}log_2\frac{1}{7})) = 0.1518$$

$$InfoGain(S,W) = 0.9403 - \frac{8}{14}(-(\frac{2}{8}log_2\frac{2}{8} + \frac{6}{8}log_2\frac{6}{8})) - \frac{6}{14} = 0.0481$$

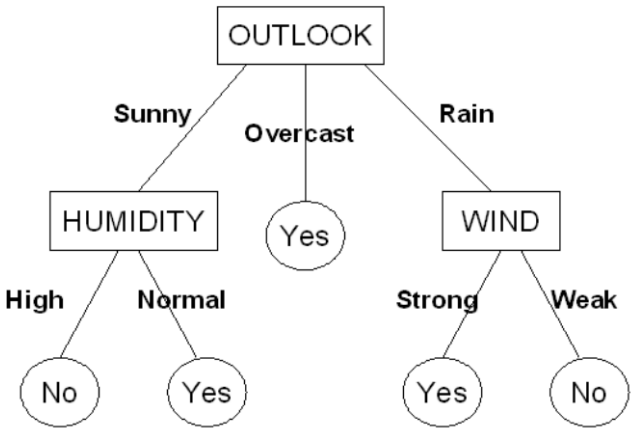
$$InfoGain(S,W) = 0.9403 - \frac{5}{14}(-(\frac{3}{5}log_2\frac{3}{5} + \frac{2}{5}log_2\frac{2}{5})) - \frac{5}{14}(-(\frac{3}{5}log_2\frac{3}{5} + \frac{2}{5}log_2\frac{2}{5}) - \frac{4}{14}0) = 0.2468$$

The first arrtribute to split on is therefore: OUTLOOK.

Next, we choose an attribute to split on in every leaf of the tree:



The fully developed tree is:



(b). 阐述如何采用剪枝策略来避免决策树的过拟合（5 分）；

比如：1. 限定每个叶子节点的最小样本数量；

2. 限制树的深度；

3. 采用 X 方策略删除信息增益小的节点；

答对一点得 3 分，答对 2 点得 5 分