Question 1. (8 points)

(a) Are the rules mutually exclusive?
Answer: No
(b) Is the rule set exhaustive?
Answer: Yes
(c) Is ordering needed for this set of rules?
Answer: Yes because a test instance may trigger more than one rule.
(d) Do you need a default class for the rule set?
Answer: No because every instance is guaranteed to trigger at least one rule.

Question 2. (10 points)

(a)
P(A = 1|+) = 3/5 = 0.6,
P(A = 1|−) = 2/5 = 0.4,
P(A = 0|+) = 2/5 = 0.4,
P(A = 0|−) = 3/5 = 0.6,
P(B = 1|+) = 1/5 = 0.2,
P(B = 1|−) = 2/5 = 0.4,
P(B = 0|+) = 4/5 = 0.8,
P(B = 0|−) = 3/5 = 0.6,
P(C = 1|+) = 4/5 = 0.8,
P(C = 1|−) = 1,
P(C = 0|+) = 1/5 = 0.2.
P(C = 0|−) = 0;
(b)
Let P (A=0, B=1, C=0) = K.

$$P(+|A=0,B=1,C=0) = \frac{P(A=0,B=1,C=0|+)\times P(+)}{P(A=0,B=1,C=0)}$$
$$= \frac{P(A=0|+)P(B=1|+)P(C=0|+)\times P(+)}{K}$$
$$= \frac{0.4\times 0.2\times 0.2\times 0{,}5}{K} = \frac{0.008}{K}$$
$$P(-|A=0,B=1,C=0) = \frac{P(A=0,B=1,C=0|-)\times P(-)}{P(A=0,B=1,C=0)}$$
$$= \frac{P(A=0|-)\times P(B=1|-)\times P(C=0|-)\times P(-)}{K} = \frac{0}{K}$$

The class label should be ‘+’;
(c)

$P(A=0|+) = (2+2)/(5+4) = 4/9,$
$P(A=0|-) = (3+2)/(5+4) = 5/9,$
$P(B=1|+) = (1+2)/(5+4) = 3/9,$
$P(B=1|-) = (2+2)/(5+4) = 4/9,$
$P(C=0|+) = (3+2)/(5+4) = 5/9,$
$P(C=0|-) = (0+2)/(5+4) = 2/9.$

(d)

Let $P(A=0, B=1, C=0) = K$

$$
\begin{aligned}
& P(+|A=0, B=1, C=0) \\
= \; & \frac{P(A=0, B=1, C=0|+) \times P(+)}{P(A=0, B=1, C=0)} \\
= \; & \frac{P(A=0|+)P(B=1|+)P(C=0|+) \times P(+)}{K} \\
= \; & \frac{(4/9) \times (3/9) \times (5/9) \times 0.5}{K} \\
= \; & 0.0412/K
\end{aligned}
$$

$$
\begin{aligned}
& P(-|A=0, B=1, C=0) \\
= \; & \frac{P(A=0, B=1, C=0|-) \times P(-)}{P(A=0, B=1, C=0)} \\
= \; & \frac{P(A=0|-) \times P(B=1|-) \times P(C=0|-) \times P(-)}{K} \\
= \; & \frac{(5/9) \times (4/9) \times (2/9) \times 0.5}{K} \\
= \; & 0.0274/K
\end{aligned}
$$

The class label should be '+'.

(e)

When one of the conditional probabilities is zero, the estimate for conditional probabilities using the m-estimate probability approach is better, since we don't want the entire expression becomes zero.

Question 3. (12 points)

(a)
P(Mileage=Hi) = 0.5
P(Air Cond=Working) = 0.625
P(Engine=Good|Mileage=Hi) = 0.5
P(Engine=Good|Mileage=Lo) = 0.75
P(Value=High|Engine=Good, Air Cond=Working) = 0.750
P(Value=High|Engine=Good, Air Cond=Broken) = 0.667
P(Value=High|Engine=Bad, Air Cond=Working) = 0.222
P(Value=High|Engine=Bad, Air Cond=Broken) = 0

(b)

$$
\begin{aligned}
& |P(Engine = Bad, Air\ Cond = Broken) \\
= \ & \sum_{\alpha\beta} P(Engine = Bad, Air\ Cond = Broken, Mileage = \alpha, Value = \beta) \\
= \ & \sum_{\alpha\beta} P(Value = \beta | Engine = Bad, Air\ Cond = Broken) \\
& \times\ P(Engine = Bad | Mileage = \alpha) P(Mileage = \alpha) P(Air\ Cond = Broken) \\
= \ & 0.1453.
\end{aligned}
$$

Question 4. (8 points)

(a)
1-nearest neighbor: +,
3-nearest neighbor: −,
5-nearest neighbor: +,
9-nearest neighbor: −.

(b)
1-nearest neighbor: +,
3-nearest neighbor: +,
5-nearest neighbor: +,
9-nearest neighbor: +.

Question 5. (12 points)

(a) Answer: 2

(b) Answer: $E = \sum_i^N (y_i - x_i^w)^2$

(c) Answer:

Use the fact that $\frac{dx^w}{dw} = x^w \log x$

$$\frac{dE}{dw} = -\sum_{i=1}^{N} 2(y_i - x_i^w)x_i^w \log x_i$$

$$= -2\sum_{i=1}^{N} \delta_i x_i^w \log x_i$$

where we write $\delta_i = (y_i - x_i^w)$

So, the required update rule is:

$$w \leftarrow w + 2\alpha \sum_{i=1}^{N} \delta_i x_i^w \log x_i$$