Q1:

(a) No. Because a record can be covered by more than one rule, for example a record can satisfy Mileage=Low and Air Conditioner=Broken at the same time.

(b) Yes. Because for every possible attribute value, we have at least one rule (Air Conditioner = Broken; Air Conditioner= Working, Engine = Good/Bad) to cover.

(c) Yes. Because the attributes have Mileage = Low together with Air Conditioner = Broken. According to the second rule the value should be High (Mileage=Low), but the fifth rule defines the value to be Low (Air Conditioner = Broken). If we do not set the order of the rules, the classifier will contain paradox.

(d) No. Because our rule-based classifier is exhaustive. Every record will trigger at least one rule and be classified. We do not need to set a default class to avoid the case that a record not triggering any rule.

Q2.

(a) P(A=1|+) = 3/5=0.6 P(A=0|+) = 0.4
P(B=1|+) = 1/5=0.2 P(B=0|+) = 0.8
P(C=1|+) = 4/5=0.8 P(C=0|+) = 0.2
P(A=1|-) = 2/5=0.4 P(A=0|-) = 0.6
P(B=1|-) = 2/5=0.4 P(B=0|-) = 0.6
P(C=1|-) = 5/5=1 P(C=0|-) = 0

(b) P(+|A=0,B=1,C=0)=

$$P(+|A=0,B=1,C=0)=\frac{P(A=0,B=1,C=0|+)P(+)}{P(A=0,B=1,C=0)}$$
$$=\frac{P(A=0|+)P(B=1|+)P(C=0|+)P(+)}{P(A=0,B=1,C=0|+)+P(A=0,B=1,C=0|-)}$$
$$=\frac{P(A=0|+)P(B=1|+)P(C=0|+)P(+)}{P(A=0|+)P(B=1|+)P(C=0|+)+P(A=0|-)P(B=1|-)P(C=0|-)}$$
$$=\frac{0.4*0.2*0.2*0.5}{0.4*0.2*0.2*0.5+0.6*0.4*0*0.5}=1$$

So the label for the test sample (A=0,B=1,C=0) is “+”.

(c) $P=\frac{n_c+mp}{n+m}$

P(A=1|+) = (3+2)/(5+4)=0.556 P(A=0|+) = 0.444
P(B=1|+) = (1+2)/(5+4)=0.333 P(B=0|+) = 0.666

P(C=1|+) = (4+2)/(5+4)=0.666   P(C=0|+) = 0.333
P(A=1|-) = (2+2)/(5+4)=0.444   P(A=0|-) = 0.556
P(B=1|-) = (2+2)/(5+4)=0.444   P(B=0|-) = 0.556
P(C=1|-) = (5+2)/(5+4)=0.778   P(C=0|-) = 0.222

(d) $P(+|A=0,B=1,C=0) = \frac{P(A=0,B=1,C=0|+)P(+)}{P(A=0,B=1,C=0)} =$

$\frac{P(A=0|+)P(B=1|+)P(C=0|+)P(+)}{P(A=0,B=1,C=0|+)+P(A=0,B=1,C=0|-)} =$

$\frac{P(A=0|+)P(B=1|+)P(C=0|+)P(+)}{P(A=0|+)P(B=1|+)P(C=0|+)+P(A=0|-)P(B=1|-)P(C=0|-)} =$

$\frac{0.444*0.333*0.333*0.5}{0.444*0.333*0.333*0.5+0.556*0.444*0.222*0.5} = 0.47324$

$$P(-|A=0,B=1,C=0) = 1 - P(+|A=0,B=1,C=0) = 1 - 0.47324 = 0.52676$$

So, P(-|A=0,B=1,C=0) is larger than P(+|A=0,B=1,C=0), so the label should be "–" given the condition (A=0,B=1,C=0)

(e) m-estimate approach is better.
Because the original approach has the problem of zero probability. (We have P(C=0|-)=0.) It is unreasonable to have P(A=0,B=1,C=0|-)=0 just because we do not observe P(C=0|-). The m-estimator can handle the zero-probability problem by adding mp to Nc and m to N.

Q3
(a)

P(Mileage)=Hi	P(Mileage)=Lo
0.5	0.5

P(Air Conditioner)=Working	P(Air Conditioner)=Broken
0.625	0.375

P(Engine | Mileage)	Engine=Good	Engine=Bad
Mileage=Hi	0.5	0.5
Mileage=Lo	0.75	0.25

P(Car Value| Air Conditioner,Engine)	Car Value=Hi	Car Value=Lo
Engine=Good	0.75	0.25

Air Conditioner=Working		
Engine=Good Air Conditioner=Broken	0.666	0.333
Engine=Bad Air Conditioner=Working	0.222	0.778
Engine=Bad Air Conditioner=Broken	0	1

(b)

$$\begin{aligned} P(Engine = Bad, AirConditioner = Broken) \\ &= P(Engine = Bad)P(AirConditioner = Broken) \\ &= \big(P(Engine = Bad|Mileage = Hi)P(Mileage = Hi) \\ &+ P(Engine = Bad|Mileage = Lo)P(Mileage = Lo)\big)P(AirConditioner \\ &= Broken) = (0.5 * 0.5 + 0.25 * 0.5) * 0.375 = 0.140625 \end{aligned}$$

Q4

(a)

1 nearest neighbor: 4.9 (+) ➔ x=5.0 classified as +

3 nearest neighbors:4.9(+),5.2(-),5.3(-)➔x=5.0 classified as –

5 nearest neighbors: 4.9(+),5.2(-),5.3(-),4.6(+),4.5(+) or 4.9(+),5.2(-),5.3(-),4.6(+),5.5(+)
➔x=5.0 classified as +

9 nearest neighbors: 4.9(+),5.2(-),5.3(-),4.6(+),4.5(+),5.5(+),3.0(-),7.0(-),0.5(-) or 4.9(+),5.2(-),5.3(-),4.6(+),4.5(+),5.5(+),3.0(-),7.0(-),9.5(-)➔x=5.0 classified as –

(b)

1-nearest neighbor: 4.9 (+)

➔ $+vote = 1 \times \frac{1}{0.1\times0.1} = 100 \qquad -vote = 0$

➔label=+

3 nearest neighbors:4.9(+),5.2(-),5.3(-)

➔ $+vote = 1 \times \frac{1}{0.1\times0.1} = 100 \qquad -vote = 1 \times \frac{1}{0.2\times0.2} + 1 \times \frac{1}{0.3\times0.3} = 36.11$

➔label=+

5 nearest neighbors: 4.9(+),5.2(-),5.3(-),4.6(+),4.5(+) or 4.9(+),5.2(-),5.3(-),4.6(+),5.5(+)

➔$+vote = 1 \times \frac{1}{0.1\times0.1} + 1 \times \frac{1}{0.4\times0.4} + 1 \times \frac{1}{0.5\times0.5} = 110.25$

- $vote = 1 \times \frac{1}{0.2\times0.2} + 1 \times \frac{1}{0.3\times0.3} = 36.11$

➔label=+

9 nearest neighbors: 4.9(+),5.2(-),5.3(-),4.6(+),4.5(+),5.5(+),3.0(-),7.0(-),0.5(-) or 4.9(+),5.2(-),5.3(-),4.6(+),4.5(+),5.5(+),3.0(-),7.0(-),9.5(-)➔x=5.0 classified as –

$\rightarrow +vote = 1 \times \frac{1}{0.1\times0.1} + 1 \times \frac{1}{0.4\times0.4} + 1 \times \frac{1}{0.5\times0.5} + 1 \times \frac{1}{0.5\times0.5} = 114.25$

$$-vote = 1 \times \frac{1}{0.2 \times 0.2} + 1 \times \frac{1}{0.3 \times 0.3} + 1 \times \frac{1}{2 \times 2} + 1 \times \frac{1}{2 \times 2} + 1 \times \frac{1}{4.5 \times 4.5} = 36.659$$

$\rightarrow$label=+

Q5.

(a) The second function

(b)

$$\hat{y} = x^w$$

$$E = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{N} (y_i - x_i^w)^2$$

(c)

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

$$w \leftarrow w - \alpha(-2w)\left(y_i x_i^{w-1} - x_i\right)$$