**Machine Learning Coursework**

Part 1

Notes:

+ : Yes

– : No

**Gini Index**

Root Node:

Root Node (14) : 7 Positive, 7 Negative

Headache+ (6) : 4 Positive, 2 Negative ; Headache– (8) : 3 Positive, 5 Negative

Spots+ (8) : 5 Positive, 3 Negative ; Spots– (6) : 2 Positive, 4 Negative

Stiff Neck+ (8) : 6 Positive, 2 Negative ; Stiff Neck– (6) : 1 Positive, 5 Negative

I(Root) = 1 – (7/14)² – (7/14)²

= 0.5

I(Headache+) = 1 – (4/6) ² – (2/6)²

= 4/9

I(Headache–) = 1 – (3/8)² – (5/8) ²

= 0.46875

I(Spots+) = 1 – (5/8)² – (3/8) ²

= 0.46875

I(Spots–) = 1 – (2/6)² – (4/6) ²

= 4/9

I(Stiff Neck+) = 1 – (6/8)² – (2/8) ²

= 0.375

I(Stiff Neck–) = 1 – (1/6)² – (5/6) ²

= 5/18

Gain(Root, Headache) = 0.5 – (6/14)(4/9) – (8/14)(0.46875)

= 0.04167

Gain(Root, Spots) = 0.5 – (8/14)(0.46875) – (6/14)(4/9)

= 0.04167

Gain(Root, Stiff Neck) = 0.5 – (8/14)(0.375) – (6/14)(5/18)

= 0.1667

**\* Split at Stiff Neck**

Stiff Neck+:

Stiff Neck+ (8) : 6 Positive, 2 Negative

Headache+ (4) : 4 Positive, 0 Negative ; Headache– (4) : 2 Positive, 2 Negative

Spots+ (5) : 4 Positive, 1 Negative ; Spots– (3) : 2 Positive, 1 Negative

I(Stiff Neck+) = 1 – (6/8)² – (2/8)²

= 0.375

I(Headache+) = 1 – (4/4) ² – 0

= 0

I(Headache–) = 1 – (2/4)² – (2/4) ²

= 0.5

I(Spots+) = 1 – (4/5)² – (1/5) ²

= 0.32

I(Spots–) = 1 – (2/3)² – (1/3) ²

= 4/9

Gain(Root, Headache) = 0.375 – 0 – (4/8)(0.5)

= 0.125

Gain(Root, Spots) = 0.375 – (5/8)(0.32) – (3/8)(4/9)

= 0.008333

**\* Stiff Neck+ split at Headache**

Stiff Neck–:

Stiff Neck– (6) : 1 Positive, 5 Negative

Headache+ (2) : 0 Positive, 2 Negative ; Headache– (4) : 1 Positive, 3 Negative

Spots+ (3) : 1 Positive, 2 Negative ; Spots– (3) : 0 Positive, 3 Negative

I(Stiff Neck–) = 1 – (1/6)² – (5/6)²

= 5/18

I(Headache+) = 1 – 0 – (2/2) ²

= 0

I(Headache–) = 1 – (1/4)² – (3/4) ²

= 0.375

I(Spots+) = 1 – (1/3)² – (2/3) ²

= 4/9

I(Spots–) = 1 – 0 – (3/3) ²

= 0

Gain(Root, Headache) = 5/18 – 0 – (4/6)(0.375)

= 0.02778

Gain(Root, Spots) = 5/18 – (3/6)(4/9) – 0

= 0.0556

**\* Stiff Neck– split at Spots**

Stiff Neck+ -> Headache+:

Headache+ (4) : 4 Positive, 0 Negative

Spots+ (2) : 2 Positive, 0 Negative; Spots– (2) : 2 Positive, 0 Negative

**\* Stop at Headache+**

**\*\* No Negative diagnosis for Headache+**

Stiff Neck+ -> Headache– Split:

Headache– (4) : 2 Positive, 2 Negative

Spots+ (3) : 2 Positive, 1 Negative ; Spots– (1) : 0 Positive, 1 Negative

I(Headache–) = 1 – (2/4)² – (2/4)²

= 0.5

I(Spots+) = 1 – (2/3)² – (1/3)²

= 4/9

I(Spots–) = 1 – 0 – (1/1)²

= 0

**\* Split at Spots and Stop**

**\*\* No more attributes**

Stiff Neck– -> Spots+:

Spots+ (3) : 1 Positive, 2 Negative

Headache– (3) : 1 Positive, 2 Negative

I(Spots+) = 1 – (1/3)² – (2/3)²

= 4/9

I(Headache–) = 1 – (1/3)² – (2/3)²

= 4/9

**\* Split at Headache and Stop**

**\*\* No more attributes**

Stiff Neck– -> Spots–:

Spots– (3) : 0 Positive, 3 Negative

Headache+ (2) : 0 Positive, 2 Negative

Headache– (1) : 0 Positive, 1 Negative

**\* Stop at Spots–**

**\*\* No Positive diagnosis for Spots–**

Diagram

Description automatically generated

**Chi-squared Statistic**

Root Node:

Root Node (14) : 7 Positive, 7 Negative

Headache+ (6) : 4 Positive, 2 Negative ; Headache– (8) : 3 Positive, 5 Negative

Spots+ (8) : 5 Positive, 3 Negative ; Spots– (6) : 2 Positive, 4 Negative

Stiff Neck+ (8) : 6 Positive, 2 Negative ; Stiff Neck– (6) : 1 Positive, 5 Negative

Headache :

Table, calendar

Description automatically generated

Spots :

Table, calendar

Description automatically generated

Stiff Neck :

Table, calendar

Description automatically generated

**\* Split at Stiff Neck**

Stiff Neck+:

Stiff Neck+ (8) : 6 Positive, 2 Negative

Headache+ (4) : 4 Positive, 0 Negative ; Headache– (4) : 2 Positive, 2 Negative

Spots+ (5) : 4 Positive, 1 Negative ; Spots– (3) : 2 Positive, 1 Negative

Headache :

Table, calendar

Description automatically generated

Spots :

Table, calendar

Description automatically generated

**\* Stiff Neck+ split at Headache**

Stiff Neck–:

Stiff Neck– (6) : 1 Positive, 5 Negative

Headache+ (2) : 0 Positive, 2 Negative ; Headache– (4) : 1 Positive, 3 Negative

Spots+ (3) : 1 Positive, 2 Negative ; Spots– (3) : 0 Positive, 3 Negative

Headache :

Table, calendar

Description automatically generated

Spots :

Calendar

Description automatically generated

**\* Stiff Neck– split at Spots**

Stiff Neck+ -> Headache+:

Headache+ (4) : 4 Positive, 0 Negative

Spots+ (2) : 2 Positive, 0 Negative; Spots– (2) : 2 Positive, 0 Negative

**\* Stop at Headache+**

**\*\* No Negative diagnosis for Headache+**

Stiff Neck+ -> Headache– Split:

Headache– (4) : 2 Positive, 2 Negative

Spots+ (3) : 2 Positive, 1 Negative ; Spots– (1) : 0 Positive, 1 Negative

**\* Split at Spots and Stop**

**\*\* No more attributes**

Stiff Neck– -> Spots+:

Spots+ (3) : 1 Positive, 2 Negative

Headache– (3) : 1 Positive, 2 Negative

I(Spots+) = 1 – (1/3)² – (2/3)²

= 4/9

I(Headache–) = 1 – (1/3)² – (2/3)²

= 4/9

**\* Split at Headache and Stop**

**\*\* No more attributes**

Stiff Neck– -> Spots–:

Spots– (3) : 0 Positive, 3 Negative

Headache+ (2) : 0 Positive, 2 Negative

Headache– (1) : 0 Positive, 1 Negative

**\* Stop at Spots–**

**\*\* No Positive diagnosis for Spots–**

Diagram

Description automatically generated