**Dataset 1 (10% missing data)**

1. Starting point (all 0.5)

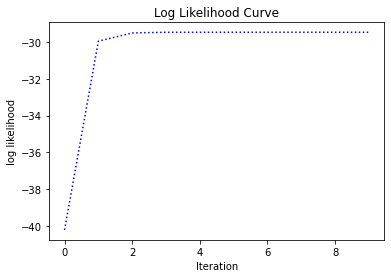
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.5 | 0.5 |
| Male | 0.5 | 0.5 |

|  |
| --- |
| P(Female) |
| 0.5 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.3656 | 0.9999 |
| Male | 0.1376 | 0.3117 |

|  |
| --- |
| P(Female) |
| 0.2735 |



1. Starting point (required in the question)

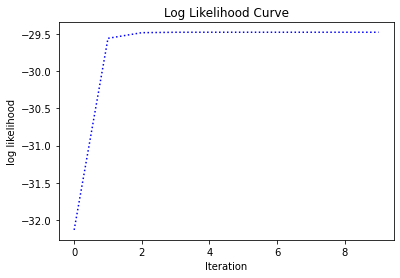
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6 | 0.7 |
| Male | 0.2 | 0.3 |

|  |
| --- |
| P(Female) |
| 0.3 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.3656 | 0.9999 |
| Male | 0.1376 | 0.3117 |

|  |
| --- |
| P(Female) |
| 0.2735 |



* **Conclusion:** Different starting points all converge to the same solution

**Dataset 2 (15% missing data)**

(Note: only 3 values are missing out of 20, hence it is not a 30% missing data but 15%)

1. Starting point (all 0.5)

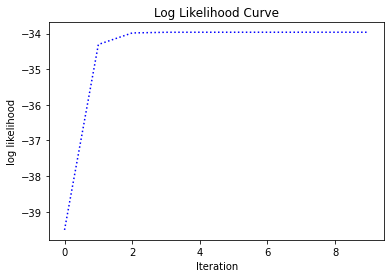
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.5 | 0.5 |
| Male | 0.5 | 0.5 |

|  |
| --- |
| P(Female) |
| 0.5 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.7655 | 0.7655 |
| Male | 0.1074 | 0.4675 |

|  |
| --- |
| P(Female) |
| 0.4445 |



1. Starting point (required in the question)

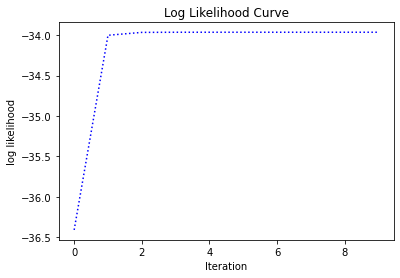
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6 | 0.7 |
| Male | 0.2 | 0.3 |

|  |
| --- |
| P(Female) |
| 0.3 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.7655 | 0.7655 |
| Male | 0.1074 | 0.4675 |

|  |
| --- |
| P(Female) |
| 0.4445 |



* **Conclusion:** Different starting points all converge to the same solution

**Dataset 3 (50% missing data)**

1. Starting point (all 0.5)

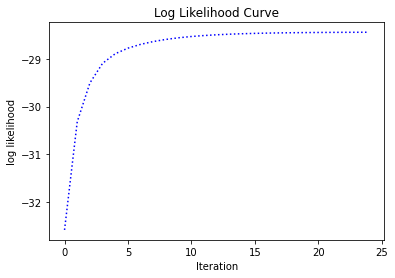
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.5 | 0.5 |
| Male | 0.5 | 0.5 |

|  |
| --- |
| P(Female) |
| 0.5 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6270 | 0.9977 |
| Male | 0.3043 | 0.1482 |

|  |
| --- |
| P(Female) |
| 0.2963 |



1. Starting point (required in the question)

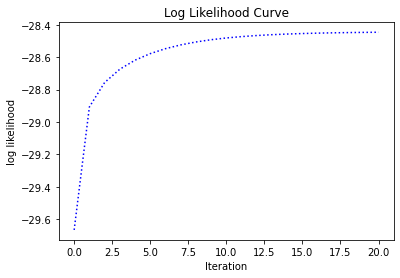
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6 | 0.7 |
| Male | 0.2 | 0.3 |

|  |
| --- |
| P(Female) |
| 0.3 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6270 | 0.9977 |
| Male | 0.3043 | 0.1482 |

|  |
| --- |
| P(Female) |
| 0.2963 |



1. Starting point

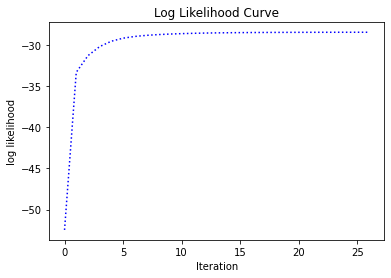
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.3 | 0.1 |
| Male | 0.9 | 0.85 |

|  |
| --- |
| P(Female) |
| 0.95 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6270 | 0.9974 |
| Male | 0.3043 | 0.1482 |

|  |
| --- |
| P(Female) |
| 0.2964 |



* **Conclusion:** Different starting points all converge to the same solution

**Dataset 4 (60% missing data)**

(Note: 12 values are missing out of 20, hence it is a 60% missing data not 70%)

1. Starting point (all 0.5)

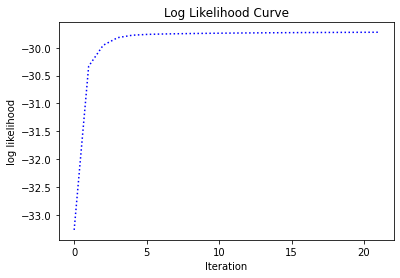
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.5 | 0.5 |
| Male | 0.5 | 0.5 |

|  |
| --- |
| P(Female) |
| 0.5 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.9115 | 0.8078 |
| Male | 0.5210 | 0.4242 |

|  |
| --- |
| P(Female) |
| 0.4583 |



1. Starting point (required in the question)

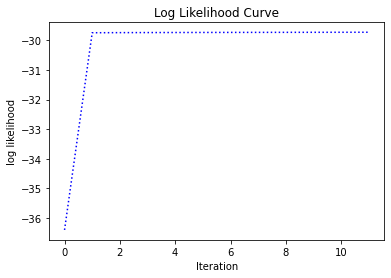
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6 | 0.7 |
| Male | 0.2 | 0.3 |

|  |
| --- |
| P(Female) |
| 0.3 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.9107 | 0.8079 |
| Male | 0.5213 | 0.4238 |

|  |
| --- |
| P(Female) |
| 0.4586 |



1. Starting point

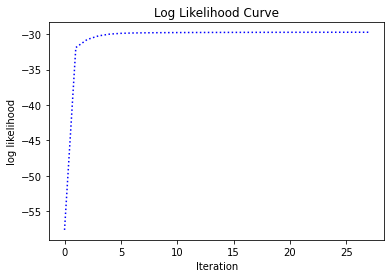
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.3 | 0.1 |
| Male | 0.9 | 0.85 |

|  |
| --- |
| P(Female) |
| 0.95 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.9101 | 0.8080 |
| Male | 0.5216 | 0.4234 |

|  |
| --- |
| P(Female) |
| 0.4590 |

****

**Conclusion:** Different starting points all converge to the same solution.

**Dataset 5 (100% missing data)**

1. Starting point (all 0.5)

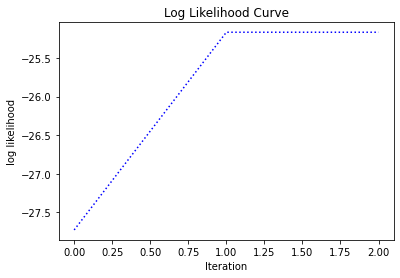
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.5 | 0.5 |
| Male | 0.5 | 0.5 |

|  |
| --- |
| P(Female) |
| 0.5 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.3 | 0.35 |
| Male | 0.3 | 0.35 |

|  |
| --- |
| P(Female) |
| 0.5 |



Final log likelihood = -25.1662

1. Starting point (required in the question)

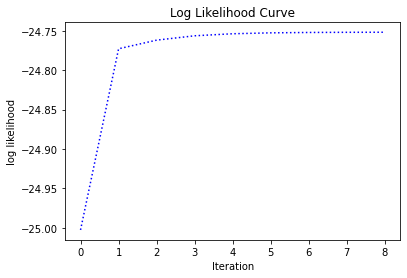
|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6 | 0.7 |
| Male | 0.2 | 0.3 |

|  |
| --- |
| P(Female) |
| 0.3 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.6388 | 0.6890 |
| Male | 0.1688 | 0.2187 |

|  |
| --- |
| P(Female) |
| 0.2792 |



Final log likelihood = -24.7520

1. Starting point

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.3 | 0.1 |
| Male | 0.9 | 0.85 |

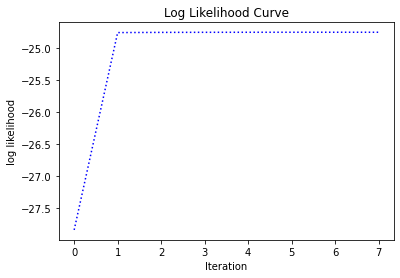
|  |
| --- |
| P(Female) |
| 0.95 |

Final Probability Tables

|  |  |  |
| --- | --- | --- |
| P(-/Gender) | Weight=1 | Height=1 |
| Female | 0.2271 | 0.2746 |
| Male | 0.8918 | 0.9622 |

|  |
| --- |
| P(Female) |
| 0.8904 |

Final log likelihood = -24.7520



* **Conclusion:** Different starting points converge to different solutions.

**Analysis**

* Do multiple starting points help in finding better solutions?

Ans: Yes. Especially in the case of 100% missing data different initializations converge to different local optima. By trying different initializations we can choose one of the solutions that give the highest log likelihood.

* Do some of the solutions have the same likelihood scores?

Ans: Yes. For example starting points 2 & 3 in 100% missing data example converged to different solutions yet gave the same likelihood score (-24.7520)

* How does the data missing rate affect your algorithm and results?

Ans: The missing rate doesn’t affect the algorithm, but it does affect the results. Lower missing rates means all initializations converge to the same solutions, but as the missing rate closes 100% the final solution becomes highly dependent on the initialization.