Linear regression

Linear regression là thuật toán dùng để xấp xỉ đầu ra một hệ thống /đối tượng theo một tổ hợp tuyến tính của các đầu vào:

Y ≈ A.X + B

***Bài 1:***

Dân số thế giới qua các năm [https://en.wikipedia.org/wiki/World\_population\_estimates](https://www.google.com/url?q=https://en.wikipedia.org/wiki/World_population_estimates&sa=D&ust=1575097242356000)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
| Population (billion) | 6.067 | 6.137 | 6.215 | 6.314 | 6.396 | 6.477 | 6.555 | 6.625 | 6.705 | 6.810 | 6.892 |

Nếu tốc độ tăng dân số thế giới không đổi theo từng năm thì quan hệ giữa dân số và năm sẽ có dạng một đường thẳng, hay:

Population = a. Year + b

***Bài 2:***

Xây dựng mô hình để ước tính lượng mỡ máu theo tuổi và cân nặng dựa vào bảng dữ liệu sau:

|  |  |  |
| --- | --- | --- |
| Weight (kilograms) | Age (Years) | Blood fat content (mg/l) |
| 84 | 46 | 354 |
| 73 | 20 | 190 |
| 65 | 52 | 405 |
| 70 | 30 | 263 |
| 76 | 57 | 451 |
| 69 | 25 | 302 |
| 63 | 28 | 288 |
| 72 | 36 | 385 |
| 79 | 57 | 402 |
| 75 | 44 | 365 |
| 27 | 24 | 209 |
| 89 | 31 | 290 |
| 65 | 52 | 346 |
| 57 | 23 | 254 |
| 59 | 60 | 395 |
| 69 | 48 | 434 |
| 60 | 34 | 220 |
| 79 | 51 | 374 |
| 75 | 50 | 308 |
| 82 | 34 | 220 |
| 59 | 46 | 311 |
| 67 | 23 | 181 |
| 85 | 37 | 274 |
| 55 | 40 | 303 |
| 63 | 30 | 244 |