**Markov Chain-Based Heading Generator**

This note explains the process of generating headings using a Markov Chain approach with an expanded corpus.

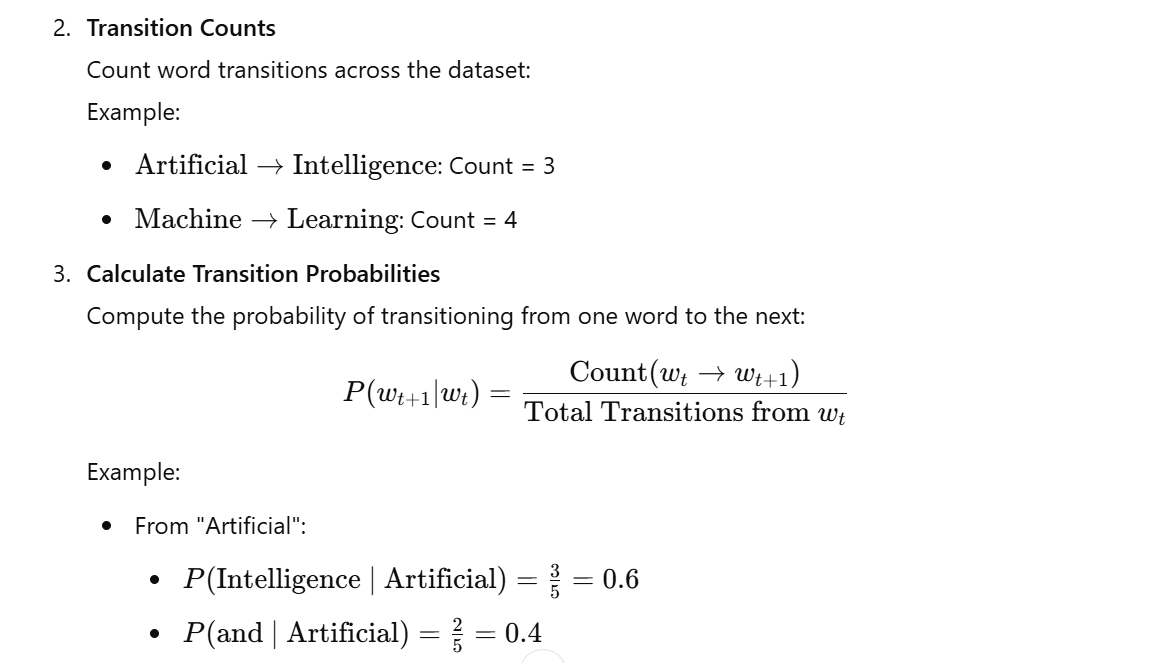
**1. Dataset (Corpus)**

We use an expanded dataset of headings:

1. "Understanding Artificial Intelligence"
2. "Introduction to Machine Learning"
3. "Artificial Intelligence and Machine Learning in Business"
4. "Applications of Deep Learning in Robotics"
5. "Machine Learning for Beginners"
6. "Deep Learning for Image Processing"
7. "The Future of Artificial Intelligence"
8. "How Machine Learning is Changing Healthcare"
9. "Artificial Intelligence in Education"
10. "The Role of Deep Learning in Natural Language Processing"
11. "Ethics in Artificial Intelligence"
12. "Machine Learning Algorithms Explained"
13. "Deep Learning Models for Autonomous Vehicles"
14. "Artificial Intelligence and the Internet of Things"
15. "Beginner's Guide to Machine Learning and AI"

**2. Step-by-Step Process**

1. **Tokenization**  
   Break each heading into sequences of words.  
   Example:
   * Heading: "Understanding Artificial Intelligence"
   * Tokens: ["Understanding", "Artificial", "Intelligence"]

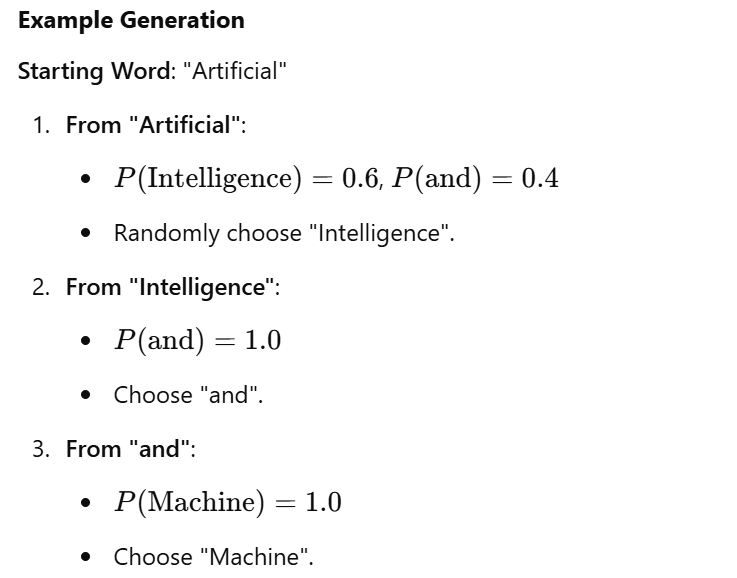


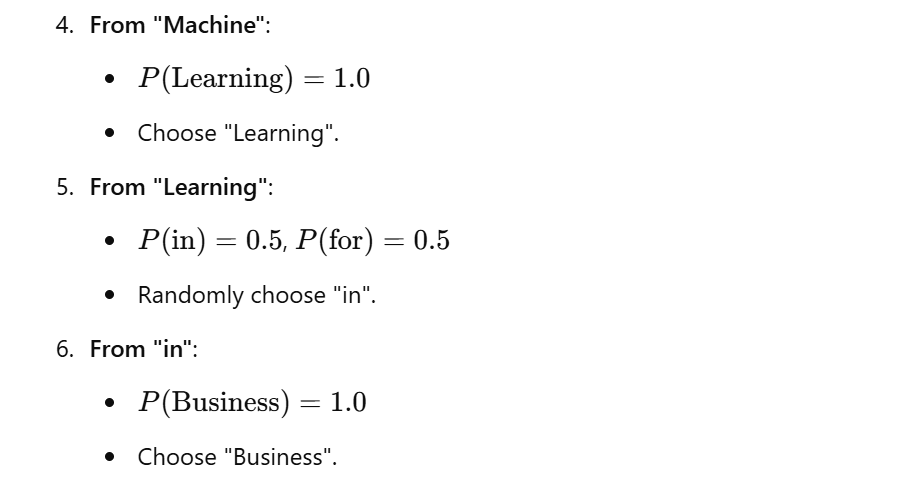
1. **Transition Matrix**  
   Represent these probabilities in a matrix:

|  |  |  |
| --- | --- | --- |
| Current Word | Next Word | Probability |
| Artificial | Intelligence | 0.60.60.6 |
| Artificial | and | 0.40.40.4 |
| Deep | Learning | 1.01.01.0 |
| Learning | in | 0.50.50.5 |
| Learning | for | 0.50.50.5 |
| Machine | Learning | 1.01.01.0 |

1. **Heading Generation**  
   Use the Markov Chain to generate a new heading:
   * Start with a random or specified word.
   * Choose the next word based on transition probabilities.
   * Repeat until reaching a terminal condition (e.g., no next word or maximum length).

**Example Generation**

****

****

**Generated Heading**:  
"Artificial Intelligence and Machine Learning in Business"

**Key Notes**

1. **Scalability**: The process scales with larger datasets.
2. **Customizability**: You can specify starting words to guide the generation process.
3. **Randomness**: Results vary due to random transitions based on probabilities.

This approach effectively generates realistic and meaningful headings based on patterns in the dataset.