# Probability Notes: Addition, Joint & Conditional Probability

## 1. Addition Rule (Union of Events)

The addition rule calculates the probability that either event A or event B occurs:  
P(A ∪ B) = P(A) + P(B) − P(A ∩ B)  
- If A and B are mutually exclusive (cannot happen together), then:  
 P(A ∩ B) = 0, so P(A ∪ B) = P(A) + P(B)  
- Used when asking: 'What is the probability that at least one of the events occurs?'

## 2. Joint Probability (Intersection of Events)

Joint probability is the probability that two events happen at the same time.

Events can be dependent or independent.

P (A ∩ B) = P(A/B). P(B) - formula for joint probability

P (A ∩ B) = Probability of A and B happening together.

P(A/B) = Conditional probability of A given B

P (B) = likelihood of happening of B

- If A and B are independent, then: then P(A/B) = P(B)  
 P(A ∩ B) = P(A) × P(B)  
- Used when asking: 'What is the probability that both events happen?'

**Note: We use conditional probability to calculate joint probability incase of dependent events.**

## 3. Conditional Probability

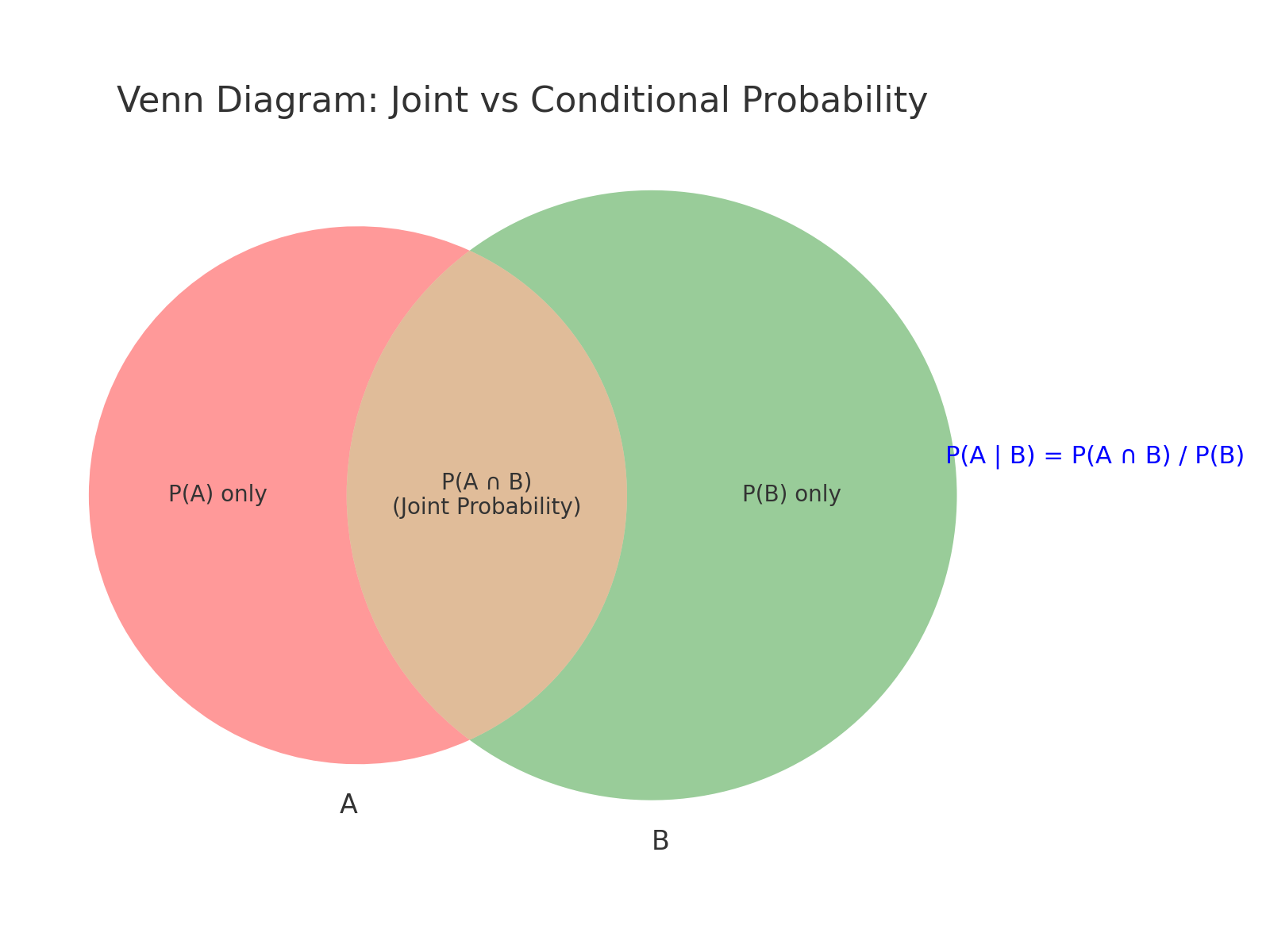
Conditional probability is the probability of one event occurring given that another has already occurred:  
P(A | B) = P(A ∩ B) / P(B), assuming P(B) ≠ 0  
- It answers: 'What is the chance of A happening, given B has already occurred?'  
- If A and B are independent, then P(A | B) = P(A)  
- Related to joint probability by:  
 P(A ∩ B) = P(A | B) × P(B)

## 4. Summary of Relationships

- Use the addition rule for calculating the probability of A OR B (union).  
- Use the multiplication rule for calculating the probability of A AND B (joint).  
- Use conditional probability when one event influences the other (dependency).  
- If A and B are independent:  
 - P(A ∩ B) = P(A) × P(B)  
 - P(A | B) = P(A)  
- If A and B are dependent:  
 - P(A ∩ B) = P(A | B) × P(B)

## 5. Venn Diagram Illustration

This diagram shows the relationship between A, B, and their intersection (joint probability).



**Summary: Where Joint Probability Appears**

| **ML Area** | **Joint Probability Role** |
| --- | --- |
| Naive Bayes | Class + all feature probabilities (jointly) |
| Bayesian Networks | Full joint via factorization |
| HMMs / Sequence Models | Joint over states and observations |
| VAEs / GANs | Model joint of data and latent space |
| Maximum Likelihood | Likelihood = joint probability of data |
| Anomaly Detection (GMMs) | Use multivariate joint distributions |
| SHAP / Explainability | Approximates joint distribution of features |