

# Recursive Retriever

Instructor

Ravi Theja

Developer Advocate Engineer, LlamalIndex



# Introduction

The real world knowledge is hierarchical

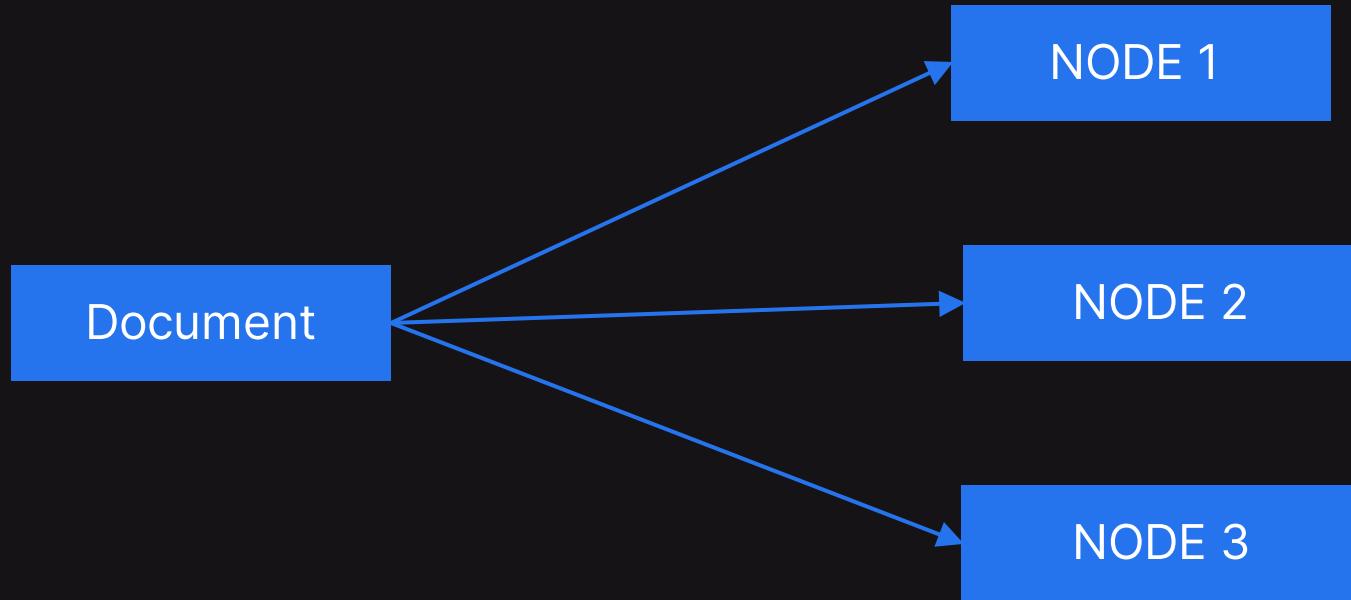
# Introduction

## Contents

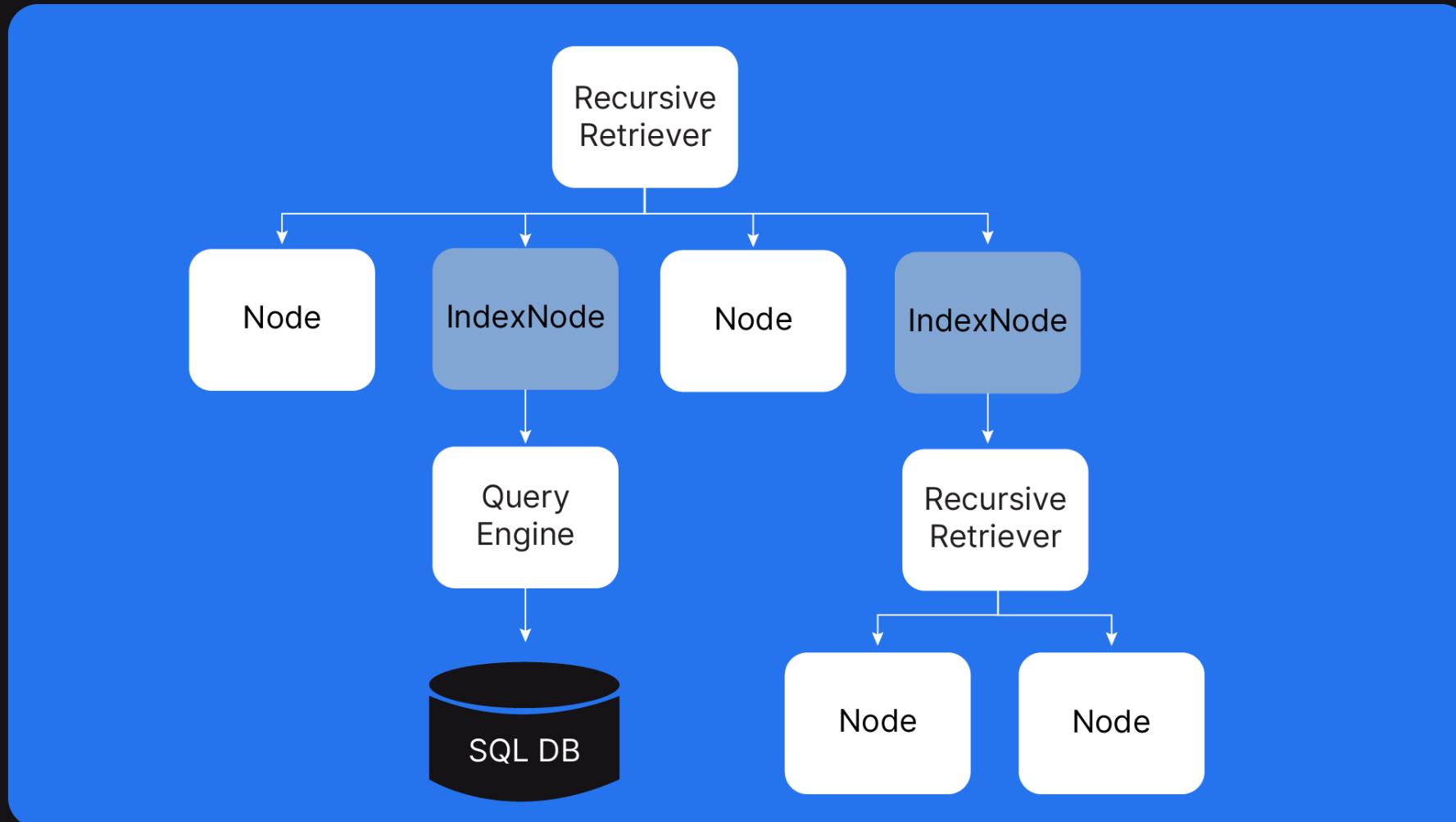
<b>Website</b>	vii
<b>Acknowledgments</b>	viii
<b>Notation</b>	xi
<b>1 Introduction</b>	1
1.1 Who Should Read This Book? . . . . .	8
1.2 Historical Trends in Deep Learning . . . . .	11
<b>I Applied Math and Machine Learning Basics</b>	29
<b>2 Linear Algebra</b>	31
2.1 Scalars, Vectors, Matrices and Tensors . . . . .	31
2.2 Multiplying Matrices and Vectors . . . . .	34
2.3 Identity and Inverse Matrices . . . . .	36
2.4 Linear Dependence and Span . . . . .	37
2.5 Norms . . . . .	39
2.6 Special Kinds of Matrices and Vectors . . . . .	40
2.7 Eigendecomposition . . . . .	42
2.8 Singular Value Decomposition . . . . .	44
2.9 The Moore-Penrose Pseudoinverse . . . . .	45
2.10 The Trace Operator . . . . .	46
2.11 The Determinant . . . . .	47
2.12 Example: Principal Components Analysis . . . . .	48

3.2 Random Variables . . . . .	56
3.3 Probability Distributions . . . . .	56
3.4 Marginal Probability . . . . .	58
3.5 Conditional Probability . . . . .	59
3.6 The Chain Rule of Conditional Probabilities . . . . .	59
3.7 Independence and Conditional Independence . . . . .	60
3.8 Expectation, Variance and Covariance . . . . .	60
3.9 Common Probability Distributions . . . . .	62
3.10 Useful Properties of Common Functions . . . . .	67
3.11 Bayes' Rule . . . . .	70
3.12 Technical Details of Continuous Variables . . . . .	71
3.13 Information Theory . . . . .	72
3.14 Structured Probabilistic Models . . . . .	75
<b>4 Numerical Computation</b>	80
4.1 Overflow and Underflow . . . . .	80
4.2 Poor Conditioning . . . . .	82
4.3 Gradient-Based Optimization . . . . .	82
4.4 Constrained Optimization . . . . .	93
4.5 Example: Linear Least Squares . . . . .	96
<b>5 Machine Learning Basics</b>	98
5.1 Learning Algorithms . . . . .	99
5.2 Capacity, Overfitting and Underfitting . . . . .	110
5.3 Hyperparameters and Validation Sets . . . . .	120
5.4 Estimators, Bias and Variance . . . . .	122
5.5 Maximum Likelihood Estimation . . . . .	131
5.6 Bayesian Statistics . . . . .	135
5.7 Supervised Learning Algorithms . . . . .	139
5.8 Unsupervised Learning Algorithms . . . . .	145
5.9 Stochastic Gradient Descent . . . . .	150
5.10 Building a Machine Learning Algorithm . . . . .	152

# Recursive Retriever



# Introduction

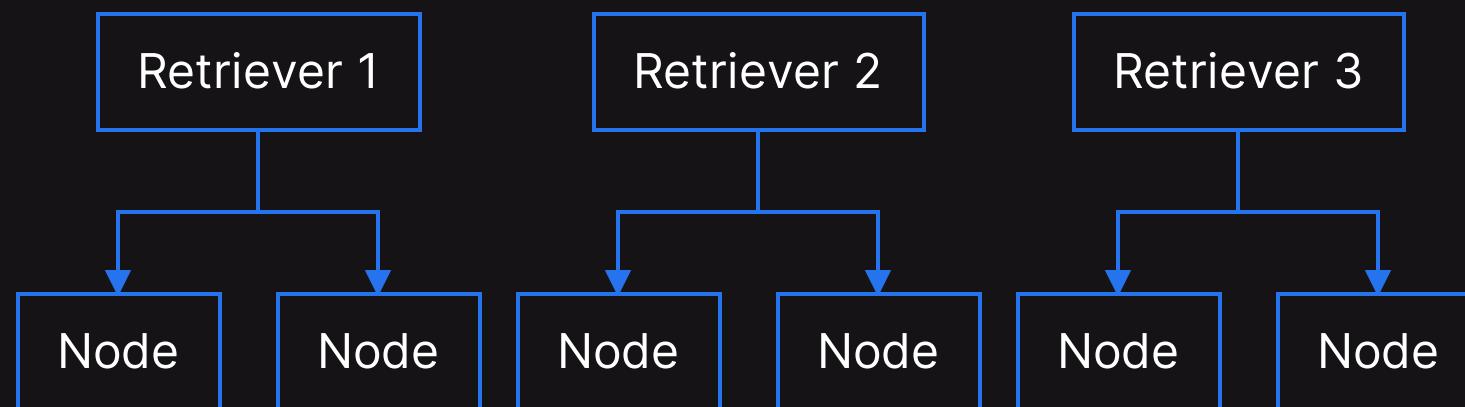


# Introduction

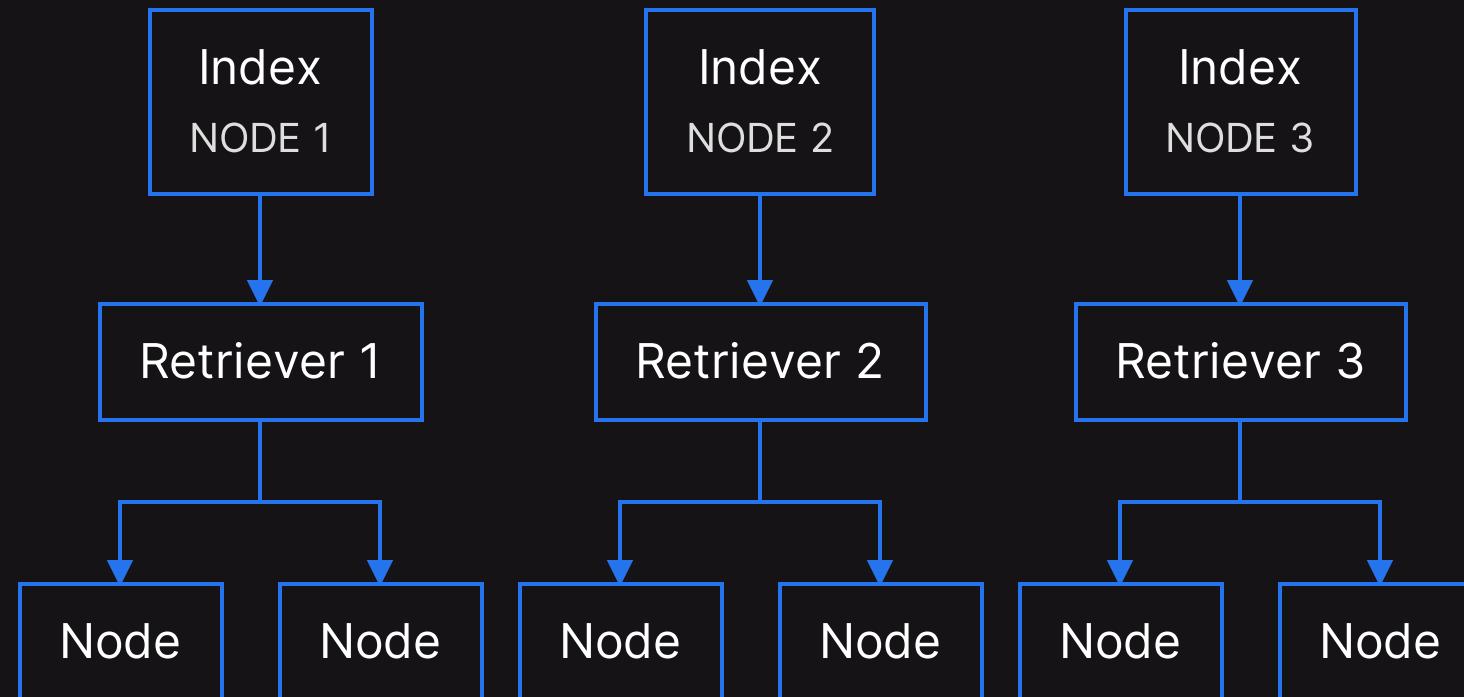
## Contents

<b>Website</b>	vii	
<b>Acknowledgments</b>	viii	
<b>Notation</b>	xi	
<b>1 Introduction</b>	1	
1.1 Who Should Read This Book?	8	
1.2 Historical Trends in Deep Learning	11	
<b>I Applied Math and Machine Learning Basics</b>	29	
<b>2 Linear Algebra</b>	31	
2.1 Scalars, Vectors, Matrices and Tensors	31	
2.2 Multiplying Matrices and Vectors	34	
2.3 Identity and Inverse Matrices	36	
2.4 Linear Dependence and Span	37	
2.5 Norms	39	
2.6 Special Kinds of Matrices and Vectors	40	
2.7 Eigendecomposition	42	
2.8 Singular Value Decomposition	44	
2.9 The Moore-Penrose Pseudoinverse	45	
2.10 The Trace Operator	46	
2.11 The Determinant	47	
2.12 Example: Principal Components Analysis	48	
3.2 Random Variables	56	
3.3 Probability Distributions	56	
3.4 Marginal Probability	58	
3.5 Conditional Probability	59	
3.6 The Chain Rule of Conditional Probabilities	59	
3.7 Independence and Conditional Independence	60	
3.8 Expectation, Variance and Covariance	60	
3.9 Common Probability Distributions	62	
3.10 Useful Properties of Common Functions	67	
3.11 Bayes' Rule	70	
3.12 Technical Details of Continuous Variables	71	
3.13 Information Theory	72	
3.14 Structured Probabilistic Models	75	
<b>4 Numerical Computation</b>	80	
4.1 Overflow and Underflow	80	
4.2 Poor Conditioning	82	
4.3 Gradient-Based Optimization	82	
4.4 Constrained Optimization	93	
4.5 Example: Linear Least Squares	96	
<b>5 Machine Learning Basics</b>	98	
5.1 Learning Algorithms	99	
5.2 Capacity, Overfitting and Underfitting	110	
5.3 Hyperparameters and Validation Sets	120	
5.4 Estimators, Bias and Variance	122	
5.5 Maximum Likelihood Estimation	131	
5.6 Bayesian Statistics	135	
5.7 Supervised Learning Algorithms	139	
5.8 Unsupervised Learning Algorithms	145	
5.9 Stochastic Gradient Descent	150	
5.10 Building a Machine Learning Algorithm	152	

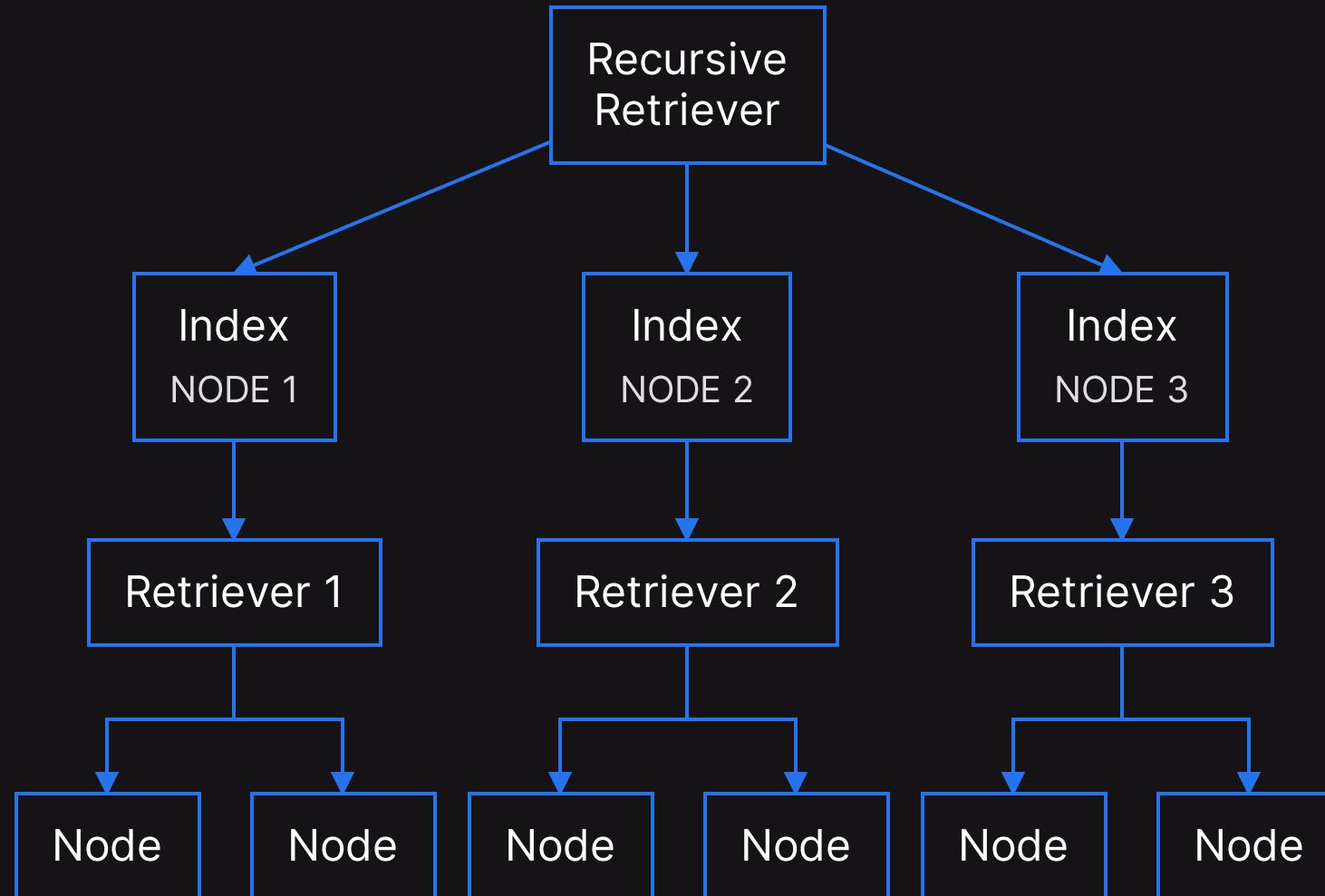
# Recursive Retriever



# Recursive Retriever



# Recursive Retriever



# Recursive Retriever

