

## Contents:

Title	Page
1. Introduction	1
2. Codon Usage Bias	3
2.1. Causes of Codon Usage Bias	4
2.1.1. Mutational biases affecting codon usage	4
2.1.2. Selection affecting codon usage	5
2.2. Measures for codon usage	7
2.2.1. Effective Number of Codons	8
2.2.2. Codon Adaptation Index	8
2.2.3. MILC-based Expression Level Predictor	9
2.2.4. Frequency of optimal codons	9
2.2.5. Codon Usage Bias	10
2.2.6. Relative codon usage bias	10
2.2.7. Codon bias index	11
2.2.8. Maximum-likelihood codon bias	12
2.2.9. Codon preference	12
2.2.10. Codon-preference bias measure	13
2.3. Predicting gene expression from CUB	14
3. ROC Curve	
3.1. Introduction	15
3.2. Area under the ROC Curve	17
4. Result and Analysis	18
4.1. Curves comparing the three methods for different organisms	20
4.2. AUC values for the different methods for each organism	23
4.3. CUB prediction in an organism with variation in gene size	23
5. Conclusion and Future works	25
6. References	26

## List of Figures:

	<b>Fig No.</b>	<b>Title</b>	<b>Page</b>
1.	2.1	GC% graph	3
2.	3.1(a)	A basic ROC graph	16
3.	3.1(b)	AUC of a discrete classifier and a probabilistic classifier	16
4.	4.1	ROC Curves of different Organisms	20
5.	4.2	Difference within an organism based on gene size	24
6.	4.3	Histogram for different gene size variation	24

## **List of Tables:**

	<b>Table No.</b>	<b>Title</b>	<b>Page</b>
1.	1.1	The Genetic Code Table	2
2.	3.1	Confusion Matrix	15
3.	4.1	List of Organisms	19
4.	4.2	Quantification of the Results	23