

**CSE 4754 Bioinformatics**

# **Tumor Sample Classification**

**TCGA: UCEC, KICH, LIHC, ESCA**



# **Team Members**

**Tasnimul Hasnat**  
**190041113**

**Nur Nasrum**  
**1900411106**

**Khaja Abdus Sami**  
**190041136**

# Introduction

## Objective

**Classification of UCEC, KICH, LIHC, ESCA  
tumor samples by tumor type using  
Logistic Regression.**



# Datasets

## Objective

**Classification of UCEC, KICH, LIHC, ESCA  
tumor samples by tumor type using  
Logistic Regression.**



# Dataset Size

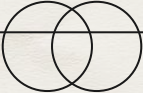
**Number of Samples: 1297**

**Number of Genes: 60660**

# Class-wise Samples (Total Count)

<b>UCEC</b>	<b>588</b>
<b>KICH</b>	<b>91</b>
<b>LIHC</b>	<b>421</b>
<b>ESCA</b>	<b>197</b>
<b>Total</b>	<b>1297</b>





# **Classification outcomes (Tumor Count)**

<b>UCEC</b>	<b>553</b>
<b>KICH</b>	<b>66</b>
<b>LIHC</b>	<b>371</b>
<b>ESCA</b>	<b>184</b>
<b>Total</b>	<b>1174</b>

# Classification and Validation Process





# Validating Classification Performance

On 260 Test Data

## Confusion Matrix

	0	1
F	23	1 <sub>1</sub>
T	1	235

# Statistical Validation

On 260 Test Data

## Classification Report

	precision	recall	f1-score	support
0	0.9583	0.9583	0.9583	24
1	0.9958	0.9958	0.9958	236
accuracy			0.9923	260



**Thank You!**