# PREDICTION OF PRIMARY TUMORS IN CANCER

**Sanushi Salgado**

A dissertation submitted in partial fulfilment of the requirement for

Bachelor of Engineering (Honours) degree in Software Engineering

**Department of Computing**

**Informatics Institute of Technology, Sri Lanka**

**in collaboration with**

**University of Westminster, UK**

**2020**

# Abstract

Machine Learning has been widely adopted in many fields, especially in healthcare. Cancer incidence and mortality are rapidly growing despite the advancements in technology.

The primary tumor of a cancer, has a significant impact on treatment options. Therefore, the diagnosis of primary tumors is vital in cancer treatment. Currently there is no efficient, accurate, less complex, cost effective solution to assist doctors in diagnosing primary tumors.

This project proposes a system that could predict the unknown primary tumors of cancer patients. The LCPN approach of Hierarchical Classification has been used for the classification process, as it gives much higher accuracy than other multiclass classification approaches and is better at handling class biasness.

Evaluation was done using 10-fold Repeated cross validation, self-evaluation, domain and technical evaluation.



**Keywords:** Machine learning, Classification, Hyperparameter tuning, Algorithm selection, Python.