**SCHOOL OF INFORMATION SCIENCES**

TERM PROJECT SYNOPSIS

|  |  |  |
| --- | --- | --- |
| Reg. No | Name | Branch |
| 171041001 | Kiran R | Computing technology and Virtualization |
| 171041003 | Harshitha R | Computing technology and Virtualization |

YouTube Video Content Analysis

|  |
| --- |
| INTRODUCTION |
| In this project, we train a classification model on the dataset Youtube-8M, which contains the largest dataset composed of over 7 million YouTube videos (450,000 hours of video) and includes video labels from a vocabulary of 4716 classes. The dataset is in tfrecords format which include video-level feature and frame-level feature, to extract and train using the dataset we use Python programming language and Tensorflow, an open-source software library for Machine learning. |
| OBJECTIVE |
| To develop classification algorithm which understands and assign video level label using Youtube-8M dataset. |
| SCOPE |
| This project consists of python code, which reads the data from Youtube-8M dataset in tfrecord format, train tensorflow model and evaluate the classification model. |
| USECASE |
| 1. Read data from tfrecord files. 2. Train classification model. 3. Evaluate the classification model. |
| TOOLS/SOFTWARE |
| Ubuntu 16.04, Tensorflow 1.3, Python 2.7. |
| HARDWARE |
| Updated in version2 of synopsis. |
| REFERENCES |
| [1] Alexandre Gauthier and Haiyu Lu,”YouTube-8M Video Classification”, Stanford University, 2017.  [2] <https://research.google.com/youtube8m/>, 14/08/17. |