

KNOWLEDGE DISTILLATION WITH SEMI- SUPERVISED MECHANISM

Talarambage Janeth Avishka Fernando

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

2021

ABSTRACT

Deep learning technologies are used in various fields including voice recognition, image classification, fraud detection, virtual assistants etc. Creating a neural network to suit the requirement is quite easy with modern frameworks like Keras. But deploying these neural networks on a less computational powered device is a challenging task.

Due to the size of neural networks, cloud technologies are used to deploy enormous neural networks. The hosted neural networks are used by a vast amount of people at the same time. These online neural networks can keep track of users' inputs. This violates the privacy of the users. Another major issue is network latency. When communicating with a remote neural network, it will take some time to get the response from the model. Availability of an active internet connection is another factor to access these neural networks.

These issues can be solved if the model is deployed on the device itself. But if the device hardware resources are minimum, these neural networks cannot be deployed on the device. Knowledge distillation can be used to compress a huge neural network and then the compressed neural network can be deployed on a less computational powered device.

Currently, knowledge distillation needs a labelled dataset to train a neural network. But labelling a dataset is an expensive task whereas it consumes time and a lot of human effort. With this research, the human effort is reduced by 90% with the new technique of semi-supervised mechanism.

Keywords: - **Neural Network, Knowledge Distillation, Semi-Supervised learning**