CSE 676: Deep Learning
Optional Project for Fall 2018

Sargur N. Srihari

University at Buffalo, The State University of New York
Buffalo, New York 14260
Contact: 716-645-6162 (O), srihari@buffalo.edu

September 19, 2018

1 Voice Classification

1.1 The Task

Classify sound data to its respective classes as explained in paper.
Use any following dataset to train and validate: Google Audioset, Urban Sounds
Achieve accuracy near to the mentioned in paper.
Understand data preprocessing, metrics used, training code and testing code
Understand the loss functions
Present the implementation and the paper

1.2 Index Terms

RNNs, LSTM, Neural Network, Convolution Neural Networks

1.3 The DL Solution

Use various Deep Learning algorithms and compare the classification accuracy.
Source code is available at https://github.com/aqibsaeed/urban-sound-classification

1.4 Project Proposal

You need to present a project proposal about four weeks into the semester. Prepare your proposal
in the form of a presentation with four parts:
1. Title (with authors), Problem Domain, description, and Data Sources
2. Variables together with their types, and proposed distributions
3. Evaluation methods.

1.5　Final Project Report

There are two deliverables:

1. project code
2. project report

The project report should describe the problem domain, data-set, algorithms used and performance (time complexity and accuracy). Use a format such as a conference paper for submission to NIPS or ICML. Include appropriate graphs and charts.