University of Toronto

Neural Networks for Machine Learning

Geoffrey Hinton - Summer 2015

Contributors: Max Smith

Latest revision: July 10, 2015

Contents

Abstract

Todo

1 Introduction

2 The Perceptron learning procedure

3 The backpropagation learning proccedure

4 Learning feature vectors for words

5 Object recognition with neural nets

5.1 Why object recognition is difficult

- Things that make it hard to recognize objects:
  - **Segmentation:** Real scenes are cluttered with other objects
    - It's hard to tell which pieces go together as parts of the same object
    - Parts of an object can be hidden behind other objects
  - **Lighting:** The intensities of the pixels are determined as much by the lighting as by the objects
  - **Deformation:** Objects can deform in a variety of non-affine ways:
    - eg. a hand-written 2 can have a large loop or just a cusp
  - **Affordances:** Object classes are often defined by how the are used
  - **Viewpoint:** Changes in viewpoint cause changes in images that standard learning methods cannot cope with

5.2 Achieving viewpoint invariance

- Several approaches:
  - Use redundant invariant features
  - Put a box around the object and use normalized pixels
  - Use replciated features with pooling. This is called "convolution neural nets"
  - Use a hierarchy of parts that have explicit poses relative to the camera

The invariant feature approach

- Extract a large, redudant set of features that are invariant under transformations
- With enough invariant features, there is only one way to assemble them into an object
- But for recognition, we must avoid forming features from parts of different objects

The judicious normalization approach

- Put a box around the object and use it as a coordinate frame for a set of normlized pixels
  - This solves the dimension-hopping problem. If we choose the box correctly, the same part of an object always occurs on the same normalized pixels
  - The box can provide invariance to many degrees of freedom: translation, rotation, scale, shear, stretch, etc.
- Chooosing the box is difficult because of: segmetnation erros, occlusion, unusual orientations
- We need to recognize the shape to get the box right (chicken and egg dilema)

The brute force normalization approach

- When training the recognizer, use well-segmented, upright images to fit the correct box
- At test time try all possble boxes in a range of positiosn and scales
  - This approach is widely used for detecting upright things like faces and house numbers in unsegmented images
  - It is much more efficient if the recognizer can cope with some variation in position and scale so that we can use a coarse grid when trying all possible boxes

6 Optimization: How to make the learning go faster

7 Recurrent neural networks

8 More recurrent neural networks

9 Ways to make neural networks generalize better

10 Combining multiple neural networks to improve generalization

11 Hopfield nets and Boltzmann machines

12 Restricted Boltzmann machines (RBMs)

13 Stacking RBMs to make Deep Belief Nets

14 Deep neural nets with generative pre-training

15 Modeling hierarchical structure with neural nets

16 Recent applications of deep neural nets