COMP0090 Intro to Deep Learning

1. **A brief history of neural networks**
2. **Mathematics refresher**
    a. Basis, orthogonal matrices, linear transformations
    b. Eigenvalues, eigenvectors, spectral decomposition (diagonalization)
    c. SVD, positive-definite matrices
    d. Taylor series, vector Taylor series
    e. Interpreting the gradient, higher derivatives
    f. logsumexp trick
3. **Optimization**
    a. Quadratic functions
    b. Gradient descent, generalised gradient update
        i. Convergence rate
        ii. Momentum, Nesterov accelerated gradient
        iii. Cons of gradient descent (coordinate system dependence, small update, etc.)
        iv. Stochastic gradient descent (SGD)
        v. Bias, variance
    c. Adaptive methods
        i. Fixed length updates
        ii. ADAM
    d. Higher order methods
        i. Newton's method, pros+cons
        ii. (Generalised) Gauss-Newton, pros+cons
        iii. Levenberg-Marquadt algorithm
4. **Automatic differentiation**
    a. Basic idea, comparison to symbolic differentiation
    b. Forward differentiation, computation tree
    c. Reverse differentiation
    d. Hessian-vector product
    e. Gauss-Newton-vector product
5. **Feedforward neural networks**
    a. Definition, diagrams
    b. Regression, comparison to standard methods
    c. Classification
        i. Logistic regression
        ii. Softmax linear regression
        iii. Loss comparison, penalty terms, transfer functions
        iv. Depth vs width
    d. Specific architectures - autoencoders, PCA
    e. NN vs computation graph
6. **Neural net initialization**
    a. Basic idea - scaling, zero-centering
    b. Decorrelating inputs
    c. Whitening

7. **Recurrent neural nets**
    a. Gradient decay/explosion
    b. Basic idea behind RNNs
    c. Mitigating memory decay
    d. LSTM, handwriting examples
    e. GRU
    f. Sequence-to-sequence models
    g. Translation examples
    h. Bidirectional RNNs
8. **Deep generative models**
    a. Basic idea, likelihood training
    b. Variational inference, VAE examples
    c. Generative adversarial networks (GANs)
    d. Image generation
9. **Practical intro to neural nets**
    a. Follow the tutorial.
10. **RNN language models**
    a. Language models
        i. Traditional approach
        ii. RNNs
    b. RNN training
        i. Vanishing/exploding gradient problems
        ii. Clipping trick for exploding gradients
        iii. Initialization and ReLUs for vanishing gradients
        iv. Other RNN implementation tricks
    c. Bidirectional RNNs, deep version
    d. Machine translation (MT)
        i. Current statistical machine translation approaches
        ii. Traditional approaches
        iii. MT using RNNs
        iv. MT using deep LSTMs
11. **Neural Nets for Image Recognition**
    a. Invariance & Assumptions
    b. Convolutions
        i. How do they work
        ii. Stride
        iii. Dilation rate
        iv. Transposed Convolutions
        v. Pooling types
    c. Architectures
        i. LeNet; AlexNet; VGG; Inception; ResNet
    d. Challenges for deep nets
        i. Batchnorm
        ii. 1x1 Convolutions