

Introduction to AI School

Levan Tsinadze - MaxinAI / Education

Welcome

Who Knows What

Math / SE

Rules

- Do not miss school
- if miss ≥ 3 : you are out
- Do not be late
- Do the homework
- Read books and papers
- Watch videos
- Etc

During The School

- Theory of ML
- Weekley Projects
- (Live) Test each 2-4 Weeks
- Project Review

Communication

- EMail (levan.tsinadze@maxinai.com)
- Slack (<http://maxinaieducation.slack.com>)
- No phone calls
- Maybe Hangout or BlueJeans calls for the paper reading and other discussions

Prepare

- Python, Numpy, SciPy, Pandas
- OpenCV, Pillow, scikit-image
- Jupyter Notebook
- Anaconda
- GitHub repository (<https://github.com>)
- Play with Colab (<https://colab.research.google.com/notebooks/welcome.ipynb>)
- Tutorial (<https://towardsdatascience.com/getting-started-with-google-colab-f2fff97f594c>)

Expectations



Reality - ML

- Find extremum of the function
- Find appropriated function
- Analyze and process data
- Write training scripts

Reality

- Ideas
- Experiments
- Rejection / Confirmation / Refinement
- New Ideas / Left or Right / New Experiments
- Mathematical Explanation
- Blog / Paper

Experiments Example

- Sequence2Sequence for Translate - > Reverse of input (Ideas Head of Sentence) Before Attention
- Super-resolution with UNet

```
self.feat_losses += [base_loss(gram_matrix(f_in), gram_matrix(f_out))*w**2 * 5e3
```

```
    for f_in, f_out, w in zip(in_feat, out_feat, self.wgts)]
```

- ResNet (Skip connections $F[n](x) := f[n](x) + x$) was result of (bruteforce) experiments
- Etc

Mathematics

- Theory
 - LA
 - Calculus
 - Probability / Statistics
 - Functional Analysis
 - Topology
 - (Applied) Category Theory
 - Etc
- Applied Mathematics
 - Probability / Statistics
 - Numerical Methods Neural ODEs
 - Etc

Software Engineering - Experiments / Deployments

- Model is Implemented as a Software
- Training is Implemented as a Software
- Inference is implemented as a Software
- Fast Translation from Math to Code

Software Engineering

- Fast and Readable Code (Algorithms and Data Structures)
- Bugless (Not easy for testing / Error After 100000...000 Iterations 10 Days)
- Testing Whatever is Possible
- DevOps (UNIX Servers, Cloud Solutions, Virtualization)
- Deployment
- Etc

Week 2

Introduction to Linear Algebra

- a) Vectors, Matrices, Tensors
- b) Linear Space, Linear Map
- c) Introduction to NumPy library
- d) Loops vs Vectorization

Week 3 / 4

Introduction to Calculus

- i) Limits of Sequences
- ii) Limits of Function
- iii) Open and Close sets
- iv) Continuous function
- v) Derivative of function
- vi) Gradient of function
- vii) Forma's theorem about function extremum and derivative

Week 5

Introduction to Probability

- a) Sigma-Algebras
- b) Probability measure
- c) Discrete vs continuous probability
- d) PDF
- e) Expectation and Variance
- f) Law of Large Numbers
- g) Distributions
- h) Metrics on distributions KL-Divergence
- i) Gaussian Mixture Models

Week 6 - 7

Introduction to Statistics

- a) Mean and Variance
- b) Central Limit Theorem
- c) Confidence intervals
- d) p-Values
- e) Hypothesis testing
- f) Regression analysis

Information Theory

- g) Information
- h) Entropy

Machine Learning Theory

- General Notions (Supervised, Unsupervised / Self-supervised / Reinforcement Learning)
- Linear Models
- Polynomial Models
- Logistic Regression / SVM
- Theory

Scikit-learn

- Scikit-learn libraries
- Off the Shelf Data
- Training
- Tuning

Data Preparation

- Normalization
- Standardization
- Feature Analysis
- Feature Engineering

Overfitting and Underfitting

- Bias vs Variance
- Hyperparameters Tuning
- Scikit-learn examples
- Experiments on Custom Data

Other Techniques

- Ensemble Methods
- Unsupervised Learning
- Boosting Libraries and Scalability
- Etc

Deep Learning

- General ANN and DNNs
- PyTorch / FastAI Libraries
- Computer Vision CNN (Classification, detection, Segmentation, Tracking, Etc)
- NLP RNN (Sentiment Analysis, NER, Translation, Etc)
 - Speech Processing
- Graphs (GCN, Embeddings, Graph Classification, Node Classification)

Projects

- Per-week
- Per-month
- Real Problems
- Real Data

Questions

