# CSC 391: Mitigating Catastrophic Forgetting

Department of Computer Science

## 1 Course Description

This independent research course explores advanced techniques in machine learning, focusing on mitigating catastrophic forgetting in artificial neural networks. Students will investigate Hebbian learning and Spike-Timing-Dependent Plasticity (STDP) inspired approaches to address this critical challenge in continual learning.

## 2 Learning Outcomes

By the end of this course, students will be able to:

- Understand the history and significance of catastrophic forgetting in neural networks
- Analyze the mathematical foundations of catastrophic forgetting
- Evaluate various approaches to mitigate catastrophic forgetting
- Implement and test Hebbian/STDP-inspired algorithms
- Conduct independent research and present findings

## 3 Course Structure

This 15-week course is divided into three main phases:

1. Foundations (Weeks 1-5)
2. Advanced Concepts and Implementation (Weeks 6-10)
3. Independent Research and Presentation (Weeks 11-15)

## 4 Weekly Schedule

### 4.1 Week 1: Introduction to Catastrophic Forgetting

#### 4.1.1 Topics

- Historical context of catastrophic forgetting
- Overview of continual learning challenges

#### 4.1.2 Readings

- McCloskey, M., & Cohen, N. J. (1989). Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. Psychology of Learning and Motivation, 24, 109-165.
- Goodfellow, I. J., Mirza, M., Xiao, D., Courville, A., & Bengio, Y. (2013). An empirical investigation of catastrophic forgetting in gradient-based neural networks. arXiv preprint arXiv:1312.6211.

#### 4.1.3 Assignment

Write a 2-page summary of the McCloskey & Cohen paper, highlighting key findings and their relevance to modern machine learning.

### 4.2 Week 2: Mathematical Foundations and Taxonomy of Mitigation Approaches

#### 4.2.1 Topics

- Mathematical formulation of catastrophic forgetting
- Overview of mitigation strategies

#### 4.2.2 Readings

- French, R. M. (1999). Catastrophic forgetting in connectionist networks. Trends in cognitive sciences, 3(4), 128-135.
- Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. Neural Networks, 113, 54-71.

#### 4.2.3 Assignment

Create a taxonomy diagram of different approaches to mitigate catastrophic forgetting, including brief descriptions of each approach.

### 4.3 Week 3: Hebbian Learning and Synaptic Plasticity

#### 4.3.1 Topics

- Principles of Hebbian learning
- Biological basis of synaptic plasticity

#### 4.3.2 Readings

- Hebb, D. O. (1949). The organization of behavior: A neuropsychological theory. New York: Wiley. (Selected chapters)

- Abbott, L. F., & Nelson, S. B. (2000). Synaptic plasticity: taming the beast. Nature neuroscience, 3(11), 1178-1183.

#### 4.3.3 Assignment

Implement a simple Hebbian learning algorithm in Python and demonstrate its basic properties.

### 4.4 Week 4: Spike-Timing-Dependent Plasticity (STDP)

#### 4.4.1 Topics

- STDP mechanisms and models
- Relevance to artificial neural networks

#### 4.4.2 Readings

- Bi, G. Q., & Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. Journal of neuroscience, 18(24), 10464-10472.
- Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nature neuroscience, 3(9), 919-926.

#### 4.4.3 Assignment

Write a report comparing and contrasting Hebbian learning and STDP, focusing on their potential applications in artificial neural networks.

### 4.5 Week 5: Review and Research Proposal

#### 4.5.1 Topics

- Review of Weeks 1-4
- Introduction to research methodologies

#### 4.5.2 Readings

- Selected papers based on individual research interests

#### 4.5.3 Assignment

Develop a research proposal (3-5 pages) outlining a potential Hebbian/STDP-inspired approach to mitigate catastrophic forgetting.

## 4.6 Week 6-7: Elastic Weight Consolidation and Synaptic Intelligence

### 4.6.1 Topics

- EWC algorithm and its variants
- Synaptic intelligence approach

### 4.6.2 Readings

- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., ... & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. Proceedings of the national academy of sciences, 114(13), 3521-3526.
- Zenke, F., Poole, B., & Ganguli, S. (2017). Continual learning through synaptic intelligence. In International Conference on Machine Learning (pp. 3987-3995). PMLR.

### 4.6.3 Assignment

Implement and compare EWC and synaptic intelligence on a simple continual learning task.

## 4.7 Week 8-9: Memory-based Approaches and Generative Replay

### 4.7.1 Topics

- Episodic memory in continual learning
- Generative replay techniques

### 4.7.2 Readings

- Lopez-Paz, D., & Ranzato, M. A. (2017). Gradient episodic memory for continual learning. Advances in neural information processing systems, 30.
- Shin, H., Lee, J. K., Kim, J., & Kim, J. (2017). Continual learning with deep generative replay. Advances in neural information processing systems, 30.

### 4.7.3 Assignment

Design and implement a simple memory-based or generative replay approach for continual learning.

## 4.8 Week 10: Hebbian/STDP-Inspired Approaches in Continual Learning

### 4.8.1 Topics

- Recent developments in Hebbian/STDP-inspired continual learning
- Challenges and opportunities

#### 4.8.2 Readings

- Aljundi, R., Babiloni, F., Elhoseiny, M., Rohrbach, M., & Tuytelaars, T. (2018). Memory aware synapses: Learning what (not) to forget. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 139-154).
- Mandge, D., & Manchanda, M. (2022). Continual learning with adaptive synapses and spiking neural networks. Frontiers in Neuroscience, 16, 855613.

#### 4.8.3 Assignment

Critically analyze a recent Hebbian/STDP-inspired approach to continual learning and propose potential improvements.

### 4.9 Week 11-13: Independent Research

#### 4.9.1 Topics

- Implementation of proposed approach
- Experimental design and evaluation

#### 4.9.2 Readings

- Customized based on individual research direction

#### 4.9.3 Assignment

Weekly progress reports and code submissions

### 4.10 Week 14: Results Analysis and Paper Writing

#### 4.10.1 Topics

- Statistical analysis of results
- Scientific writing and visualization

#### 4.10.2 Readings

- Guidelines on scientific writing and result presentation

#### 4.10.3 Assignment

Draft of research paper (6-8 pages, conference paper format)

### 4.11 Week 15: Final Presentation and Submission

#### 4.11.1 Topics

- Effective research presentation
- Peer review process

## 5 Grading

The student is expected to dedicate at least 12 hours per week towards work for this course, including during the finals period. All work for the course will be completed by December 17, 2024.

| **Component** | **Weight** |
|---|---|
| Weekly Assignments and Participation | 30% |
| Research Proposal | 15% |
| Implementation and Experiments | 30% |
| Final Paper | 25% |

## 6 Resources

- TensorFlow or PyTorch for neural network implementation
- Jupyter Notebooks for interactive development and visualization
- GitHub for version control and code sharing
- Overleaf for collaborative LaTeX writing

## 7 Office Hours

By appointment, to be scheduled via email.

*Note: This syllabus is subject to change based on the progress and interests of the student. Any modifications will be discussed and agreed upon by both the instructor and the student.*