



Online-Appendix

„Predicting Stock Market Trends Using Convolutional Neural Networks: A Deep Learning Approach“

Ege Özkul

Technical University of Munich

Junior Management Science 11(1) (2026) 195-226

Appendix

| Figure A.1: Out-of-Sample Performance of Vision Transformer Architectures with Varying Hyperparameters | | | | | | |
|---|--------------------------|-----------|-------------|-------------------|-------------------|-------------------|
| | Basic Vision Transformer | | | Swin Transformer | | |
| Dataset | 5D/60R | 5D/60R | 5D/60R | 5D/60R | 5D/60R | 5D/60R |
| Learning Rate | 1.00E-04 | 1.00E-04 | 1.00E-04 | 1.00E-04 | 1.00E-06 | 1.00E-04 |
| Label Smoothing | Not applied | Applied | Not applied | Applied | Applied | Applied |
| Early Stopping | 10 epochs | 10 epochs | 10 epochs | No early stopping | No early stopping | No early stopping |
| Patch Size | 4 x 4 | 4 x 4 | 4 x 4 | 4 x 4 | 4 x 4 | 4 x 4 |
| H-L Return | 0.06 | 0.06 | 0.06 | 0.07 | 0.07 | 0.03 |
| H-L Sharpe Ratio | 0.93 | 0.87 | 0.78 | 1.14 | 1.14 | 0.37 |

Description: The table presents the out-of-sample performance of different Vision Transformer architectures, comparing Basic Vision Transformer and Swin Transformer under varying hyperparameters. The models are evaluated on the I5R60 dataset, with different settings for learning rate, label smoothing, and early stopping. The results include H-L returns and H-L Sharpe ratios

| Figure A.2: Out-of-Sample Performance of CaiT with Varying Hyperparameters | | | | | | |
|---|-------------------|-------------|----------|----------|----------|----------|
| Dataset | 5D/60R | 5D/60R | 5D/5R | 20D/5R | 20D/5R | 20D/5R |
| Learning Rate | 1.00E-05 | 1.00E-04 | 1.00E-05 | 1.00E-04 | 1.00E-04 | 1.00E-05 |
| Label Smoothing | Not applied | Not applied | Applied | Applied | Applied | Applied |
| Early Stopping | No early stopping | 5 epochs | 5 epochs | 5 epochs | 5 epochs | 5 epochs |
| Patch Size | 4 x 4 | 4 x 4 | 4 x 4 | 4 x 4 | 4 x 4 | 8 x 8 |
| H-L Return | 0.05 | 0.08 | 0.80 | 0.74 | 0.74 | 0.75 |
| H-L Sharpe Ratio | 0.68 | 1.17 | 7.01 | 6.69 | 6.69 | 6.98 |

Description: The table presents the out-of-sample performance of different CaiT architecture under varying hyperparameters. The models are evaluated on the I5R60 dataset, with different settings for learning rate, label smoothing, and early stopping. The results include H-L returns and H-L Sharpe ratios