TIME SERIES ANALYSIS AND FORECASTING

RESOURCES WEEK 2

PART 2: MACHINE LEARNING FOR TIME SERIES

2.1 MACHINE LEARNING AND TIME SERIES

CORE TEXTBOOKS

Book	Chapter / Section	Focus
Manu Joseph – *Modern Time Series Forecasting with Python* (2nd ed., 2024)	Part 2 *Machine Learning for Time Series* (Feature Engineering chapters)	Lag features, rolling windows, Fourier terms, calendar-based features
Géron, A. – *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed., 2019)	Ch. 15 *Processing Sequences Using RNNs and CNNs*	Handling sequential/time features for ML models
Brownlee, J. – *Machine Learning Mastery with Python* (2016)	Ch. 9 *Feature Engineering for Time Series*	Practical guide to lag features, rolling statistics, Fourier transforms
Esling, P. & Agon, C. – *Time-Series Data Mining* (2012)	Survey sections on feature extraction	Broader ML feature engineering concepts applied to time series

RESEARCH PAPERS

Author(s) & Year	Title	Focus
Chen, Y., et al. (2019)	*Probabilistic Forecasting with Temporal Convolutional Neural Network*	Temporal CNNs for time-dependent features
Zhou, H., et al. (2022)	*FEDformer: Frequency Enhanced Decomposed Transformer for Long-term Series Forecasting*	Fourier-based feature decomposition for ML forecasting
Montero-Manso, P., et al. (2020)	*FFORMA: Feature-based Forecast Model Averaging*	Feature engineering to select forecasting models

TUTORIALS & GUIDES

Resource	Focus
Analytics Vidhya – 6 Powerful Feature Engineering Techniques for Time Series **(2019)**	Lag features, rolling windows, date/time features
Medium – Advanced Feature Engineering for Time Series Data **(2024)**	Fourier terms, decomposition-based features
GeeksforGeeks – Feature Engineering for Time Series Data **(2024)**	Statistical, time-domain, frequency-domain features

2.2 PREPARING THE TARGET

CORE TEXTBOOKS

Book	Chapter / Section	Focus
Hyndman & Athanasopoulos – *Forecasting: Principles and Practice* (3rd ed., 2021)	*Section 3.6 "Transformations and adjustments" & Section 9.1 "Stationarity and differencing"*	Log / Box-Cox transforms; differencing; decomposition via STL; stationarity vs non-stationarity.
Manu Joseph – *Modern Time Series Forecasting with Python* (2nd ed., 2024)	Chapters on *Target Preparation / Preprocessing* (often covering transforms, differencing, decomposition)	Likely covers Box-Cox, STL, detecting trends / seasonality in applied Python context *(exact chapter depends on edition)*
Box, Jenkins, Reinsel & Ljung – *Time Series Analysis: Forecasting and Control*	Early chapters on stationarity, differencing, transformations	The classic treatment of trend vs non-stationarity and how differencing / transformations are used.
Shumway & Stoffer – *Time Series Analysis and Its Applications* (4th ed.)	Parts covering stationarity, unit roots, decomposition, variance stabilizing transformations	Gives rigorous statistical framework for why these are needed, plus how to test / decompose.

RESEARCH PAPERS

Author(s) & Year	Title	Focus
Nwogu, E. C. & Ibebuogu, C. C. (2016)	*Some Tests for Seasonality in Time Series Data*	Discusses Kruskal-Wallis, Friedman, Edwards tests for detecting seasonality (after detrending).

Lelwala, E. I., Seamasinghe, W. M., & Gunarathna, K. M. L. M. (2024)	*Nonparametric Approach to Detecting Seasonality in Time Series: Application of the Kruskal-Wallis (KW) Test on Tourist Arrivals to Sri Lanka*	Empirical example: detrend + test seasonality via Kruskal-Wallis.
Hu, Z. et al. (2020)	*Modified Mann-Kendall Trend Test for Hydrological Time Series*	Trend detection in presence of autocorrelation etc. Good for illustrating practical constraints of trend tests.
Chen, S., Ghadami, A., Epureanu, B. I. (2020)	*Practical Guide of Using Kendall's τ in the Context of Forecasting Critical Transitions*	Covers how Kendall's tau (and Mann-Kendall) are used; issues like window size, autocorrelation; good reference for trend detection.

TUTORIALS & GUIDES

Resource	Focus
"Stationarity and Differencing" in *Forecasting: Principles & Practice* (fpp2 / fpp3)	Explains concept of stationarity, how differencing & transformations (log/Box-Cox) help, and how to visually & statistically check.
"Transformations and Adjustments" section in *Forecasting: P&P*	How to choose and apply Box-Cox / log transforms; use of STL decomposition to separate trend/seasonal/noise.
"Mastering Stationarity in Time Series" blog article	Intuitive guide: what stationarity means; how trends / seasonality / variance non-constancy show up, how to use transformations & differencing. (codefinity.com)
"KB Time-Series Data Part 3: Making a time-series stationary" (Medium)	Shows log / Box-Cox; visual & decomposition methods; applying STL etc. (Medium)

2.3 MACHINE LEARNING MODELS

CORE TEXTBOOKS

Book	Chapter / Section	Focus
Manu Joseph – *Modern Time Series Forecasting with Python* (2nd ed., 2024)	Chapters on *Machine Learning Models / Tree-based Methods & Gradient Boosting*	Implementation of trees, random forests, gradient boosting in time series settings; trade-offs

Géron, A. – *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed., 2019)	Chapters on decision trees, ensemble methods, boosting	Strong ML theory + practice; good for understanding strengths & limitations of each model type
Hastie, T., Tibshirani, R., & Friedman, J. – *The Elements of Statistical Learning* (2nd ed., 2009)	Chapters on Trees, Bagging, Random Forests, Boosting (e.g. Ch. 8-10)	Deep theory + empirical behaviour; what underlies random forest vs boosting
James, G., Witten, D., Hastie, T., Tibshirani, R. – *An Introduction to Statistical Learning (ISLR)* (2nd ed., 2021)	Chapter 8 *Tree-based Methods*	More accessible intro to trees / random forests / boosting, with R / Python examples

RESEARCH PAPERS

Author(s) & Year	Title	Focus
Bandara, K., Bergmeir, C., & Smyl, S. (2020)	"Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach"	Uses tree-based and ML methods alongside others, shows trade-offs, benchmarking
Hyndman, R. J. & Athanasopoulos, G. (2018)	"A Brief History of Machine Learning for Time Series Forecasting"	Survey/comparison of linear vs tree-based vs statistical methods; good background on strengths & weaknesses
Snap et al. (2022)	"LightGBM vs XGBoost for Retail Demand Forecasting"	Practical case study comparing boosted trees in forecasting context

TUTORIALS & PRACTICAL GUIDES

Tutorial / Guide	What it Covers
Time Series Forecasting with Python & Machine Learning (Decision Trees, Regressions & Gradient Boosting) **— Rahul Mohan (2023)**	Uses a retail sales dataset; shows how to build features, use decision trees, gradient boosting (XGBoost) for forecasting; includes code examples. Good for getting hands-on.
Forecasting with Decision Trees and Random Forests **— "Sarem Seitz" blog (2022)**	Demonstrates autoregressive random forest; shows how to generate forecast intervals by sampling individual decision trees; includes plots and practical implementation.

Time Series Forecasting with XGBoost and LightGBM **— mLearning.ai**	Focus on ensemble tree-based models (XGBoost, LightGBM) for forecasting energy consumption; shows lag features and rolling statistics, plus feature engineering. Very good for comparing different boosted tree libraries.
Comparing Transform Techniques for Tree-Based Models **— Snowflake Engineering Blog (2024)**	How to transform target or feature data to help tree-based models extrapolate better; experiments with different transformations; caution about extrapolation. Useful to show trade-offs.

2.4 COMBINING MODELS

CORE TEXTBOOKS

Book	**Chapter / Section**	Focus
Woodward, W. A., Sadler, B. P., & Robertson, S. – *Time Series for Data Science: Analysis and Forecasting* (1st ed., 2022)	The book *discusses techniques of ensemble modelling for combining information from several strategies* — ensembles appear throughout those model-comparison / model selection / model assessment sections.	A fairly recent, accessible book; it covers classical and ML methods plus ensembles in forecasting in a way meant for data science audiences; good pedagogical style.
(Classic forecasting textbooks like) — Box, Jenkins, Reinsel & Ljung — *Time Series Analysis: Forecasting and Control*	In later chapters, discussions of combining forecasts are touched on (though not always as a full dedicated chapter) via topics like model averaging or combining ARIMA forecasts.	Useful for historical / theoretical context of combining forecasts.
Hyndman & Athanasopoulos – *Forecasting: Principles and Practice*	Sections on model comparison / multiple forecasting methods can segue into combining methods. (They don't have a full "ensembles chapter" but they show comparisons that motivate combinations.)	Useful for learning about benchmarks and how combinations often outperform single models; good source of examples.

RESEARCH PAPERS

Authors & Year	**Title**	**Focus**
Gastinger, J. (2021)	*A Study on Ensemble Learning for Time Series Forecasting*	Evaluates a variety of ensembling methods: weighted averages, stacking, ensemble selection; shows how stacking meta-models can be built with different base learners (linear, random forest, XGBoost) and how ensemble weights or meta-models can improve over simple combinations.

Adhikari & Agrawal (2013)	*Combining Multiple Time Series Models Through A Robust Weighted Mechanism*	Proposes a robust non-linear weighted ensemble, considering correlation among forecasts from individual models; compared with simple linear combinations, shows better performance. Good to illustrate trade-offs in weighting vs simplicity.
Cawood & van Zyl (2021)	*Feature-weighted Stacking for Nonseasonal Time Series Forecasts: A Case Study of the COVID-19 Epidemic Curves*	Combines statistical and ML models (Prophet, LSTM etc.) using stacking; introduces meta-features indicating forecast accuracy; demonstrates improved forecasts over base models and simple blending. Matches content: stacking, blending, when to use.
Cawood & van Zyl (2022)	*Evaluating State of the Art, Forecasting Ensembles- and Meta-learning Strategies for Model Fusion*	Empirical comparison of many ensembling strategies (averaging, stacking, hybrid) over large time series datasets (M4 etc.); good to show that choice depends on data, resources, and goals.
Lu, M. et al. (2023)	*A Stacking Ensemble Model of Various Machine Learning …*	For hydrological prediction (runoff); uses RF, AdaBoost, XGB as base models, an attention-based meta model; shows how stacking can improve over simple averaging or individual models. Good practical example.

TUTORIALS & PRACTICAL GUIDES

Guide / Tutorial	What it Covers
Stacking Ensemble Machine Learning With Python **— Jason Brownlee (2021)**	Tutorial on stacking (meta-learner), combining multiple base learners; shows regression & classification stacking using scikit-learn; good for implementing stacking / blending.
Stacking Ensemble Models to improve Forecasting **— Skforecast user guide**	Example of stacking regressors (base models + final estimator) in time series forecasting; shows how to build stacked ensembles in forecasting context.
Ensemble Averaging & Weighted Ensembles with Modeltime Ensemble (R) **— R-bloggers post**	Introduces modeltime.ensemble package, showing how to do average, median, weighted ensembles and stacking in R for time series data; helpful for demonstrating simple to advanced combining methods.
Medium – *Boosting, Stacking and Bagging for Ensemble Models for Time Series* **— Kyle Jones (2025)**	High-level guide to combining methods; shows Python code examples of bagging, stacking etc. Helps illustrate trade-offs of simple vs more complex ensembles. (Medium)

PART 3: DEEP LEARNING FOR TIME SERIES FORECASTING

3.1 USING DEEP LEARNING FOR FORECASTING

CORE TEXTBOOKS

Title	Chapter / Section	Focus
***Deep Learning in Time Series Analysis* by Arash Gharehbaghi (2024) – CRC Press**	Early chapters introduce MLPs, feed-forward networks, CNNs/RNNs in time series contexts.	A newer book focussed specifically on applying deep learning to time series.
***Deep Learning for Time Series Cookbook* – O'Reilly**	Recipes using PyTorch; sections for building basic MLPs and perhaps CNNs for forecasting.	Very practical; helps translate theory into working code.
Deep time series forecasting: a survey" – many authors, 2025	Covers general paradigms, including feed-forward and convolutional-based models, feature extraction etc.	Good for getting a map of the landscape: what DL architectures are being used, when & why.

RESEARCH PAPERS

Author(s) & Year	Title	Focus
Kong et al. (2025)	*Deep Learning for Time Series Forecasting: A Survey*	Summarizes architectures, including simple feed-forward networks / MLPs and CNNs; gives strengths/weaknesses; useful to show why MLPs are simple but limited.
Chen, Si-An et al. (2023)	*TSMixer: An All-MLP Architecture for Time Series Forecasting*	Introduces TSMixer: a model built from MLPs (no recurrence) that mixes information in time & across features. Good example of how far feed-forward / MLP-based methods can go.
Kaushik, Jain, Jain, Dash (2019)	*Performance evaluation of deep neural networks for forecasting time-series with multiple structural breaks and high volatility*	Compares MLP vs CNN vs RNN/LSTM/GRU on financial time series; gives empirical insight on when MLPs work vs when more complex sequence models are needed.

3.2 RECURRENT NEURAL NETWORKS

CORE TEXTBOOKS

Book	Chapter / Section	Focus

***Deep Learning* by Ian Goodfellow, Yoshua Bengio, and Aaron Courville**	Chapter "Sequence Modeling: Recurrent and Recursive Nets" (especially sections on LSTM, GRU, vanishing/exploding gradients)	The theory bible. Rigorous foundation. Great for understanding why RNNs work the way they do, not just coding.
***Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (Aurélien Géron, 2nd ed., 2019)**	Chapter on "Processing Sequences, Time Series, and Natural Language"	It covers RNN / LSTM / GRU, benefits & trade-offs. Very good for practitioners, includes code.
***Modern Time Series Forecasting with Python* by Manu Joseph (2nd Edition)**	The chapter(s) on Deep Learning Models cover RNN / LSTM / GRU as applied in forecasting tasks	Useful as a "bridge" between theory and applied work.
Deep Learning with Python by François Chollet (2nd ed., 2021, Manning)	Chapters on sequence modeling (LSTMs, GRUs). Examples in Keras; includes text and time series forecasting.	Very accessible, written by the creator of Keras. Hands-on + conceptual clarity.

RESEARCH PAPERS

Author(s) & Year	Title	Focus
Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2019)	*A Comparative Analysis of Forecasting Financial Time Series Using ARIMA, LSTM and BiLSTM*	Compares ARIMA vs LSTM vs BiLSTM; evidence that BiLSTM often outperforms LSTM and ARIMA.
Azari, A., Papapetrou, P., Denic, S., & Peters, G. (2019)	*Cellular Traffic Prediction and Classification: a comparative evaluation of LSTM and ARIMA*	Shows LSTM outperforms ARIMA when series long and features adequate; good example.
ArunKumar et al., 2022	Comparative Analysis of Gated Recurrent Units (GRU) and LSTM	Empirical comparison of LSTM vs GRU vs simpler RNNs; reports what settings / data length favor GRU or LSTM.
Siami-Namini, S., & Siami Namin, A. (2018)	*Forecasting Economics and Financial Time Series: ARIMA vs LSTM*	LSTM significantly reduces error vs ARIMA in many cases; supports your point about when LSTM outperforms simpler/linear/statistical baselines.

TUTORIALS & PRACTICAL GUIDES

Tutorial / Guide	What it covers
Machine Learning Mastery –** Time Series Prediction with LSTM **Recurrent Neural Networks in Python with Keras	How to build an LSTM model for forecasting a classical time series (Airline Passengers), including data shaping, windowing, handling long sequences etc.; good baseline + comparison.
"*Time Series Forecasting Using GRU: A Step-by-Step Guide***" (Saravanan, ~2024)**	Builds a GRU model; shows how GRU can be simpler/faster than LSTM; includes implementation in Keras/TensorFlow. (Medium)
"*Learn by Example: RNN / LSTM / GRU Time Series***" – Kaggle notebook**	Hands-on notebook comparing RNN, LSTM, GRU; shows differences in training time, memory, predictions; useful for demonstration.
***Codecademy –** RNN PyTorch Time Series Tutorial: Complete Guide*	Builds RNNs / LSTMs in PyTorch; good for illustrating sequence memory, gating, etc.

3.3 CONVOLUTIONAL NETWORKS

CORE TEXTBOOKS

Book	Chapter / Section	Focus
***Deep Learning* by Ian Goodfellow, Yoshua Bengio, and Aaron Courville**	Chapter 9: *Convolutional Networks*. Covers CNNs, dilated convolutions, and their application in sequence modeling.	Strong theoretical foundation; helps explain why CNNs and dilated convolutions are effective for time series forecasting.
***Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (Aurélien Géron, 2nd ed., 2019)**	Chapter on *Processing Sequences, Time Series, and Natural Language* → covers CNNs, 1D convolutions, and their application in time series forecasting.	Practical guide with code examples; useful for implementing CNNs in forecasting tasks.
***Modern Time Series Forecasting with Python* by Manu Joseph (2nd Edition)**	The chapter(s) on Deep Learning Models → should cover CNNs and TCNs as applied in forecasting tasks.	Directly applicable to time series forecasting; includes code implementations.

RESEARCH PAPERS

Author(s) & Year	Title	What it contributes re CNN / TCN in forecasting
Bai, S., Zhan, X., & Kolter, J. Z. (2018)	*An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling*	Introduces TCNs with dilated convolutions and residual blocks; compares TCNs with RNNs and LSTMs.
Lea, C., Vidal, R., Reiter, A., & Hager, G. D. (2016)	*Temporal Convolutional Networks for Action Segmentation and Detection*	Proposes TCNs for sequence modeling; demonstrates their effectiveness over RNNs in action segmentation tasks.
Bai, S., Kolter, J. Z., & Koltun, V. (2018)	*Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling*	Provides empirical comparisons of TCNs with RNNs and LSTMs; highlights advantages of TCNs in sequence modeling tasks.

TUTORIALS & PRACTICAL GUIDES

Tutorial / Guide	What it covers
Temporal Convolutional Networks for Time Series Analysis **(Medium)**	Implements a TCN for time series forecasting using PyTorch; covers data preparation, model building, and evaluation.
Temporal Convolutional Networks for Time Series Forecasting in Python **(Medium)**	Introduces TCNs and their application in time series forecasting; provides code examples using the Darts library.
PyTorch Temporal Convolutional Networks **(Kaggle)**	Provides a simple baseline implementation of TCNs in PyTorch; suitable for beginners.

3.4 FORECASTING WITH TRANSFORMERS

CORE TEXTBOOKS

Book	Chapter / Section	Focus
Deep Learning — Ian Goodfellow, Yoshua Bengio, Aaron Courville (2016, MIT Press)	Chapter 10: *Sequence Modeling: Recurrent and Recursive Nets*	Provides foundational understanding of sequence modeling, including the theory behind attention mechanisms. Note: Transformers are not in the original edition, but the theory of sequences and attention is covered; good for conceptual grounding.

Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow — Aurélien Géron (2nd ed., 2019)	Chapter on *Processing Sequences, Time Series, and Natural Language*	Introduces sequence modeling and attention mechanisms; includes some modern adaptations (code in Keras/TensorFlow).
Modern Time Series Forecasting with Python — Manu Joseph (2nd Edition, 2023)	Deep Learning chapters	Covers time series forecasting with modern neural networks, including transformer-based architectures like Autoformer and Informer.
Deep Learning for Time Series Forecasting — Arash Gharehbaghi (2024, CRC Press)	Chapters on Transformer architectures for time series	Focused on deep learning applications to time series, including transformers (self-attention, TCN hybrids), causal masking, and forecasting strategies. Good for bridging theory and implementation.

RESEARCH PAPERS

Author(s) & Year	**Title**	Focus
Nayak, G.H.H. (2024)	*Transformer-based deep learning architecture for time series forecasting*	Discusses the application of transformer-based architectures in time series forecasting, highlighting their advantages in capturing long-range dependencies.
Caetano, R. (2025)	*Transformer-Based Models for Probabilistic Time Series Forecasting with Explanatory Variables*	Evaluates various transformer-based models, including Vanilla Transformer, Informer, Autoformer, ETSformer, NSTransformer, and Reformer, for probabilistic time series forecasting in retail.
Su, L. (2025)	*A systematic review for transformer-based long-term series forecasting*	Provides a comprehensive review of transformer-based models for long-term time series forecasting, discussing their effectiveness and limitations.

TUTORIALS & PRACTICAL GUIDES

Tutorial / Guide	**Focus**
How to make a Transformer for Time Series Forecasting with PyTorch **(Max Halford, Medium)**	Step-by-step implementation of a Transformer model for time series forecasting using PyTorch, including data prep and training.
Transformers for Time Series Forecasting in Python **(Unit8 Darts library blog)**	Documentation for using the Darts library (PyTorch backend) to implement Transformer models for forecasting, with code examples.

Temporal Fusion Transformer for forecast (DARTS) ***(Kaggle notebook from komalnalawade)***	Practical Python notebook that applies a transformer variant / TFT-style model via Darts for forecasting tasks (runnable on Kaggle / Colab)
Transformers in Time-series Analysis: A Tutorial ***(Ahmed et al., 2022)***	An arXiv tutorial with code pointers and worked examples; includes Python references and links to implementations. Useful for conceptual background + practical pointers.

3.5 SPECIALIZED ARCHITECTURES

CORE TEXTBOOKS

Book	Relevant Sections / Coverage	Why it's useful
Modern Time Series Forecasting with Python — Manu Joseph (2nd Ed., 2023)	Deep learning forecasting chapters — includes coverage of N-BEATS, TFT, and transformer-based models like Autoformer/Informer.	Good match for specialized transformer-based architectures; very recent, close to state-of-the-art.
Deep Learning in Time Series Analysis — Arash Gharehbaghi (2023, CRC Press)	Chapters on advanced architectures (Transformers, TFT, hybrid models).	Academic style, but covers specialized models; good for conceptual depth.
Time Series Forecasting Using Generative AI — Banglore Vijay Kumar Vishwas & Sri Ram Macharla (2025)	Covers Time Series Transformer, Temporal Fusion Transformer, Transformer models, WaveNet, and more "generative AI" models applied to forecasting.	Strong focus on DeepAR and probabilistic forecasting; complements TFT/N-BEATS.
Advanced Forecasting with Python — Joos Korstanje (2021)	As before: includes Python code and covers DeepAR, Prophet, LSTMs; not quite N-BEATS etc., but solid for specialized models like DeepAR + transformer-style models.	Retains DeepAR and other advanced models; also Python examples; good complement.

RESEARCH PAPERS

Author(s) & Year	Title	Focus
Oreshkin, B. N., Carpov, D., Chapados, N., Bengio, Y. (2019)	*N-BEATS: Neural Basis Expansion Analysis for Interpretable Time Series Forecasting*	Introduces N-BEATS, a deep learning model using stacks of fully connected residual blocks; good interpretability; strong performance on M3, M4, and Tourism datasets. (arXiv)
Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021)	*Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting*	Defines TFT; handles past & known future inputs, static metadata, attention, gating; strong trade-off between accuracy & interpretability.

TUTORIALS & PRACTICAL GUIDES

Resource	Focus
Temporal Fusion Transformer: Demand Forecasting with PyTorch-Forecasting	Demonstrates using TemporalFusionTransformer (Python) on the Stallion dataset; shows how TFT handles real/future known covariates.
nixtlaverse / NeuralForecast: NBEATS model usage	Python usage of N-BEATS via the NeuralForecast (or Nixtla) framework; shows how to fit, predict, use interpretable variant etc.
Temporal Fusion Transformer in PyTorch ***(Kaggle notebook by tomwarrens)***	Python notebook with TFT implemented via PyTorch or PyTorch-Forecasting; good for seeing hands-on use.