Master Seminar WS 18/19
Landnutzungsklassifikation - als Wettbewerb

Patrick Schäfer

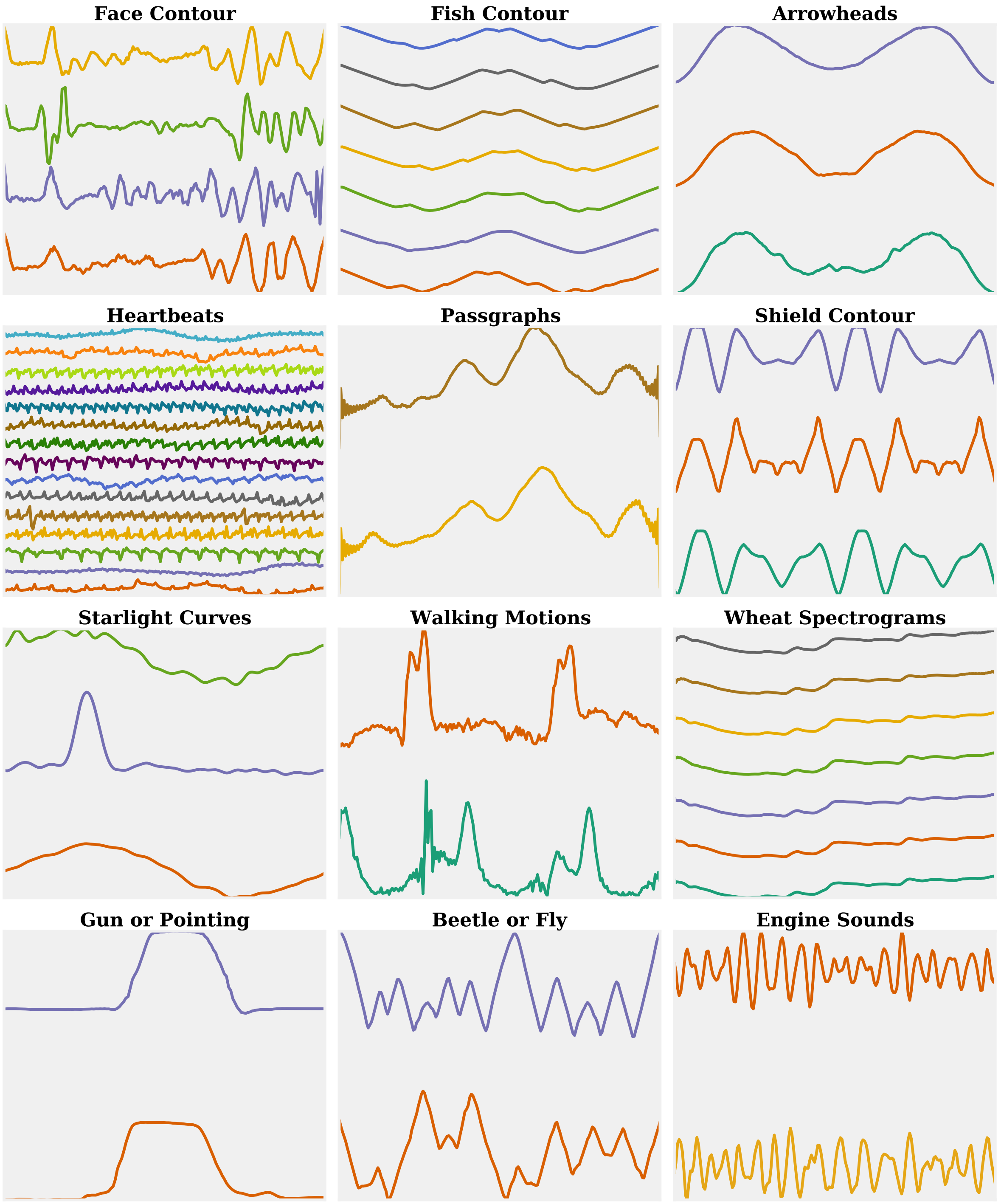
URL: https://hu.berlin/landnutzung

ToC

- Today:
  - Today 26.11.: Introduction and Group Formation
- Next Friday:
  - By next Friday 2.11.: Choice of Topic
  - Next Friday 2.11.: Primer on Time Series Analytics

Motivation

- **Temporal data** is common in many data mining applications
- Application domains range from:
  - Sensor data: environmental sensors measure temperature, pressure humidity
  - Medical devices: electrocardiogram (ECG) and electroencephalogram (EEG)
  - Financial market: stock prices, economic indicators, product sales
  - Meteorological data: sediments from drill holes, **earth observation satellite data**
- See UCR time series archive for sample datasets

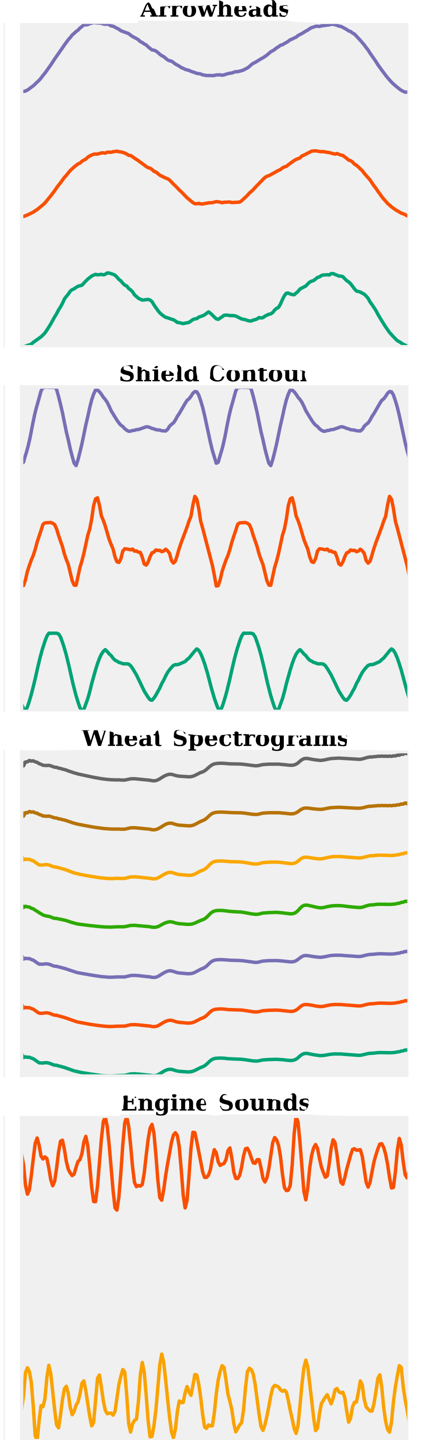


Time Series Definition

- Definition: A Time Series is a sequence (ordered collection) of n real values at time stamps $(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_{\boldsymbol{n}})$:

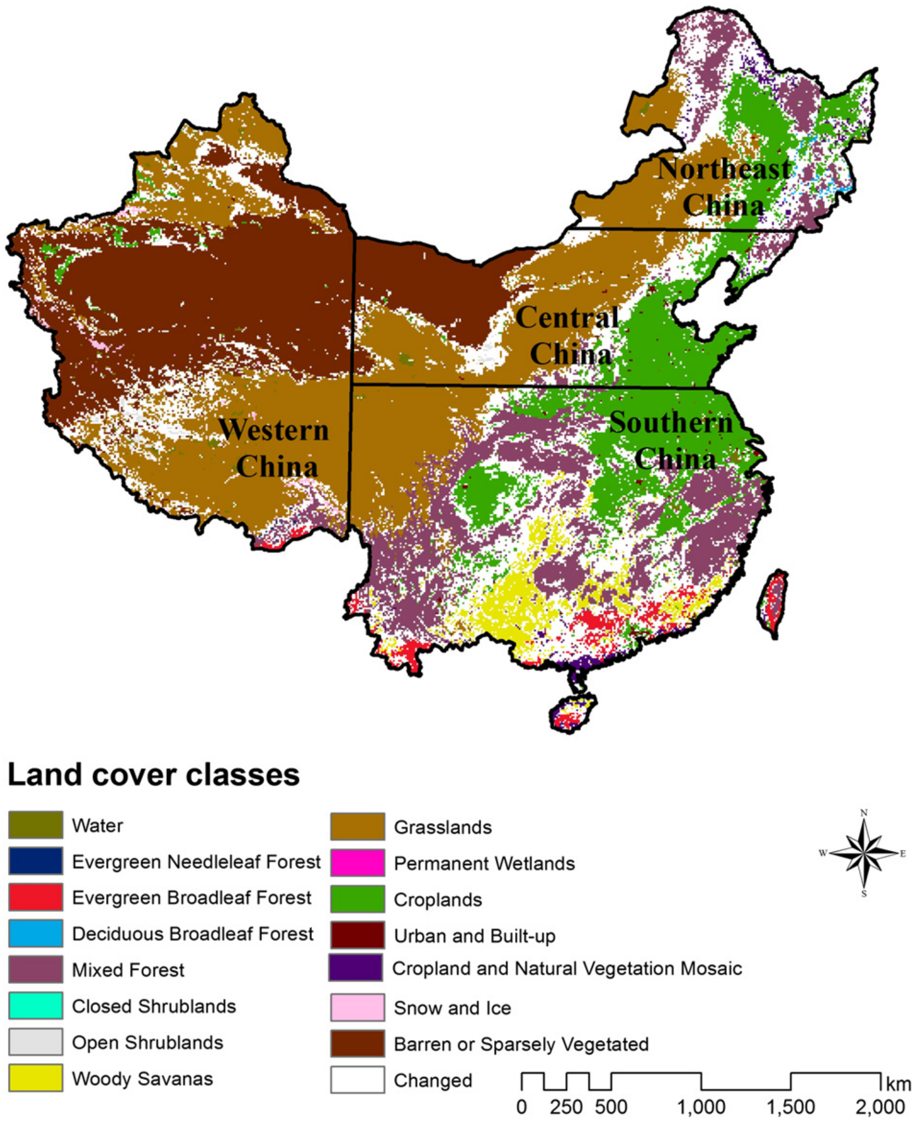
$$\boldsymbol{T} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$$

- Time Series may be univariate or multivariate
  - Univariate: a single value y_i is associated with each time stamp t_i.
  - Multivariate: m values $y_i = (k_1, \ldots k_m)$ are associated with each time stamp t_i.
- The dimensionality of a time series refers to the number of values at each time stamp



Land Cover Mapping

- Land cover is the physical material that covers the surface of the earth like grass, asphalt, urbanized areas, trees, bare soil, water, …
- The are two primary methods for capturing information on land cover:
  - Small-scale: Field survey
  - Large-scale: Remote sensed imaginary from satellites
- Land cover changes relate to (natural) processes:
  - Disasters: flooding, forest disturbance and degradation, wildfire
  - Anthropogenic activities: urbanization, agriculture, deforestation
  - Climate change

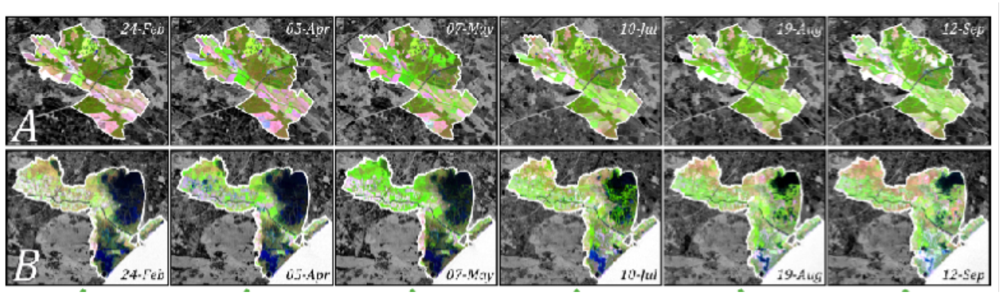


He, Yaqian, Eungul Lee, and Timothy A. Warner. "A time series of annual land use and land cover maps of China from 1982 to 2013 generated using AVHRR GIMMS NDVI3g data." *Remote Sensing of Environment* 199 (2017): 201-217.

Satellite-based Earth Observation

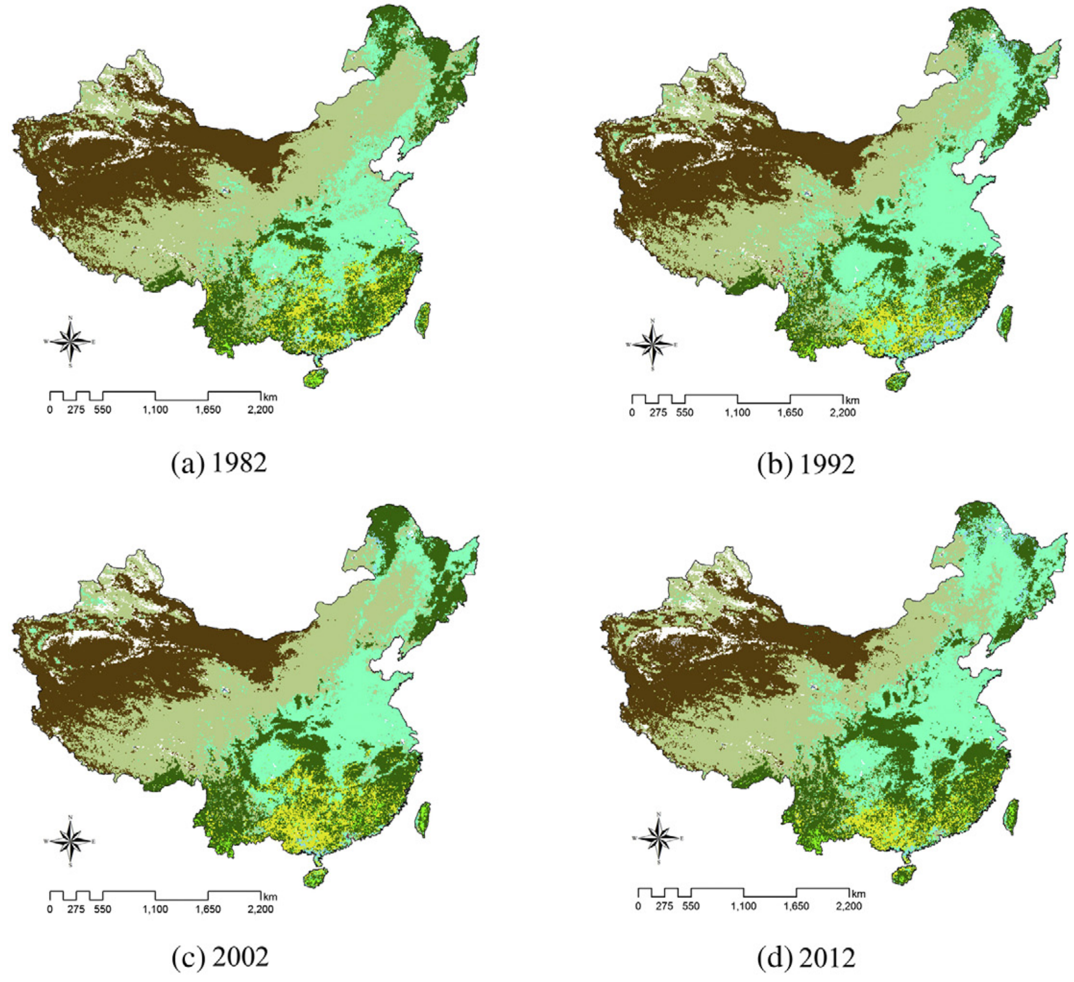
- Satellites periodically observe Earth's surface and acquire large, temporal, multi-spectral image sets
  - American Landsat 8, 30 m spatial resolution, 8 spectral bands, every 16 days
  - European Sentinel-2, 10 to 20 m spatial resolution, every 5 days

- Satellites enable the identification of the nature of spatial-temporal changes from space



„Satellite Image Time Series Analysis by RNNs –Preliminary Results" – CES Det. Changement, France, 2017

Change of annual land cover of china



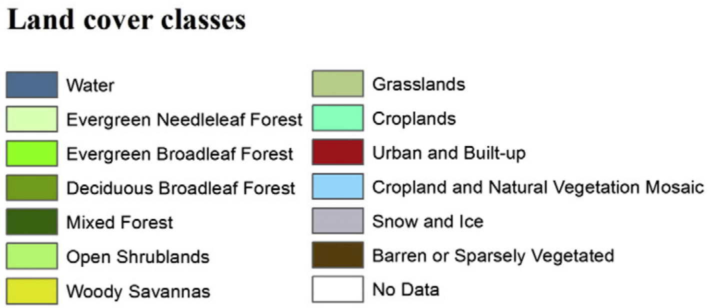
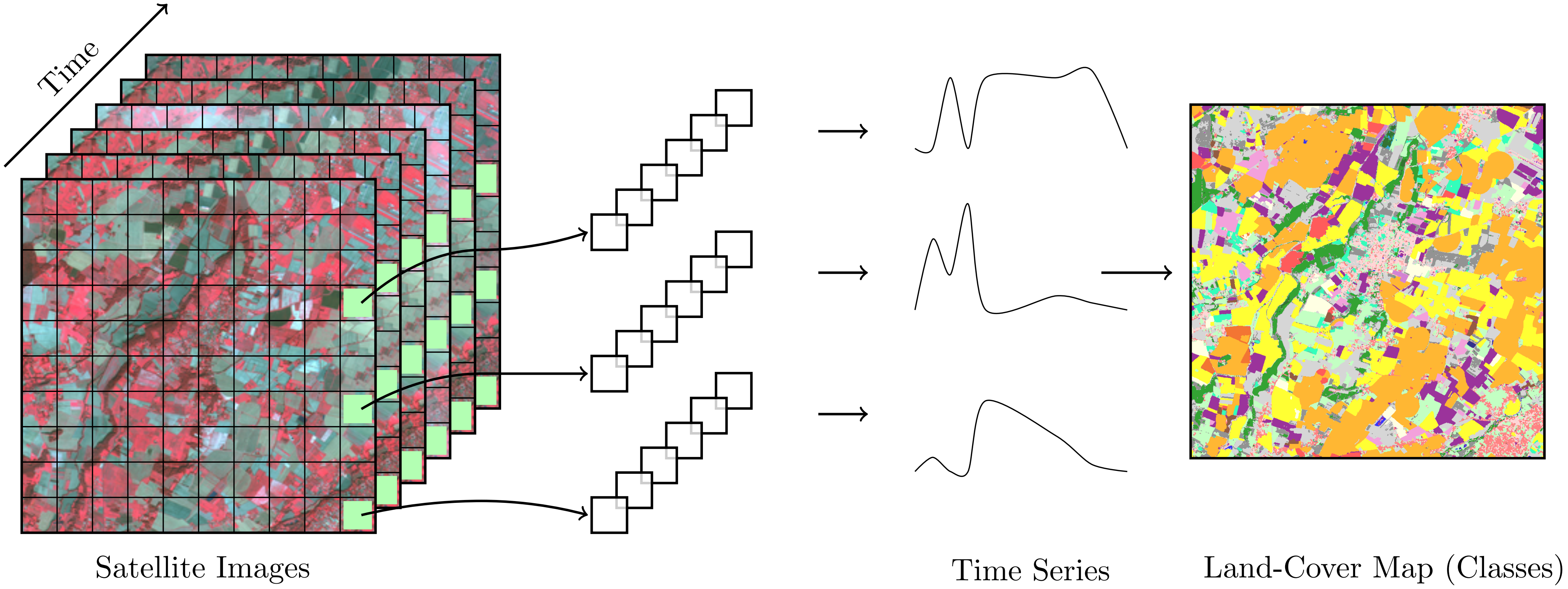


Fig. 12. Annual LULC maps of China, produced by random forest classification. (a) 1982, (b) 1992, (c) 2002, and (d) 2012.

He, Yaqian, Eungul Lee, and Timothy A. Warner. "A time series of annual land use and land cover maps of China from 1982 to 2013 generated using AVHRR GIMMS NDVI3g data." *Remote Sensing of Environment* 199 (2017): 201-217.

From satellite images to pixel time series



Tan, Chang Wei, Geoffrey I. Webb, and François Petitjean. "Indexing and classifying gigabytes of time series under time warping." *Proceedings of the 2017 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, 2017.

A smoothed NDVI time-series

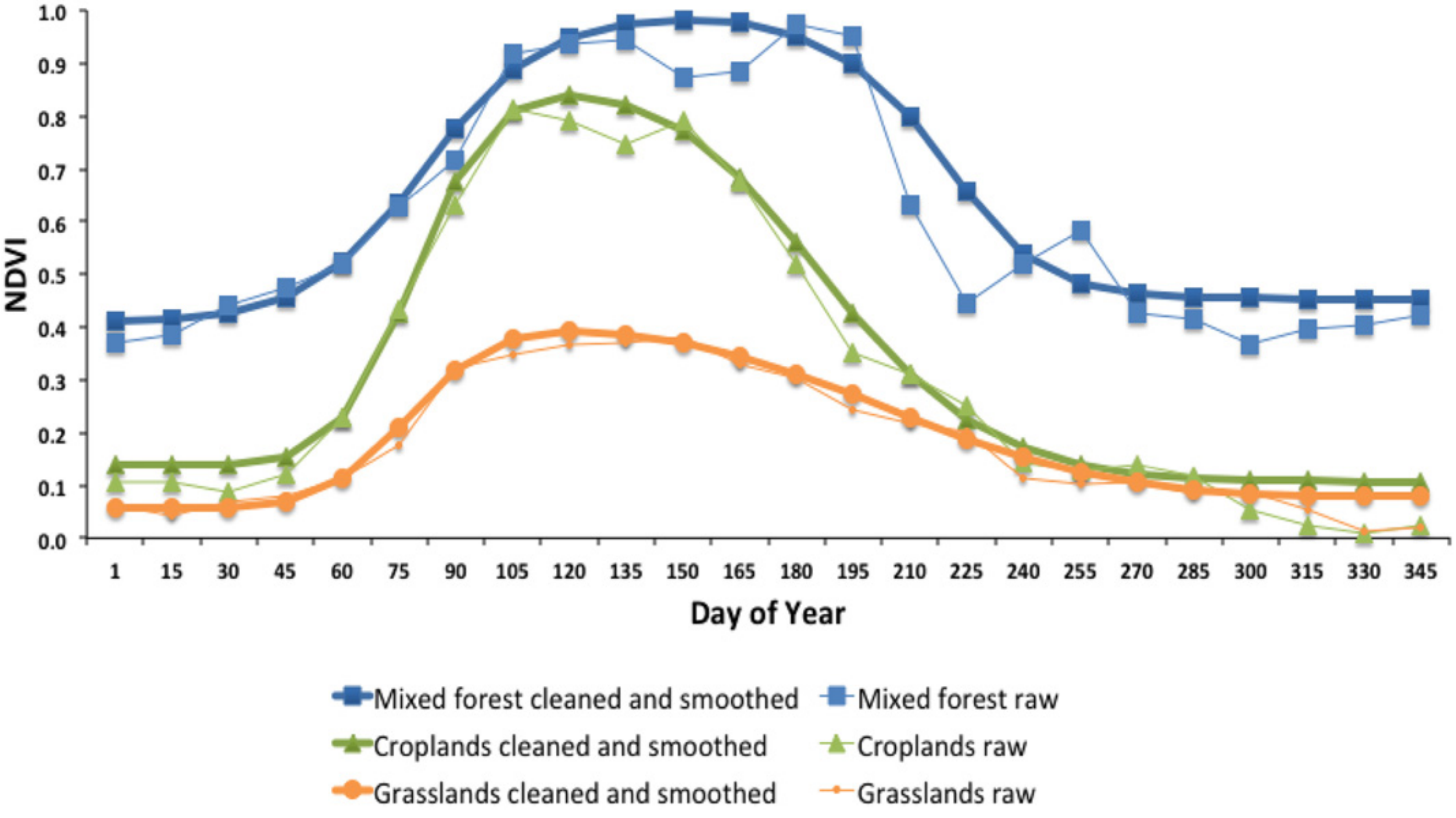


Fig. 3. Raw, and cleaned and smoothed NDVI time-series of mixed forest, croplands, and grasslands.

He, Yaqian, Eungul Lee, and Timothy A. Warner. "A time series of annual land use and land cover maps of China from 1982 to 2013 generated using AVHRR GIMMS NDVI3g data." *Remote Sensing of Environment* 199 (2017): 201-217.

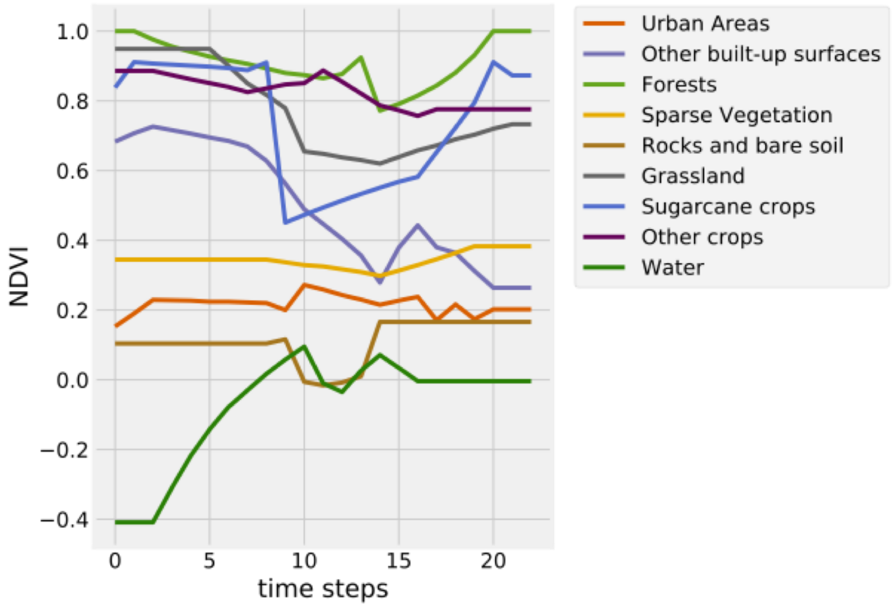
Why time series?

- We could also work with single images (single date)
- Large gaps: In some areas, the number of observations is low (due to shadows, clouds or haze):
- Thus, repeated observations of the same location increase the chance of cloud-free observations
- Repeated observations allow for identifying temporal trends (flooding, wildfire, anthropogenic activities (urbanization, agriculture), forest disturbance and degradation)
- Time series analysis has shown to be superior (more accurate) when compared to single-date methods

maps.google.de

Spectral Sensors / Features

- Land surfaces absorb and reflect sunlight differently
- Satellite are equipped with multi-spectral sensors
  - Visible spectrum:
    Blue, green, red (absorbed by green vegetation)
  - Near-infrared (absorbed by water)
  - Short-wave infrared
- Derived Features: 3 computed indices:
  - for vegetation NDVI,
  - for water NDWI,
  - for brightness BI
- This allows **machine-assisted mapping** of land cover from space



A pixel time series of the Normalized Difference Vegetation Index (NDVI)

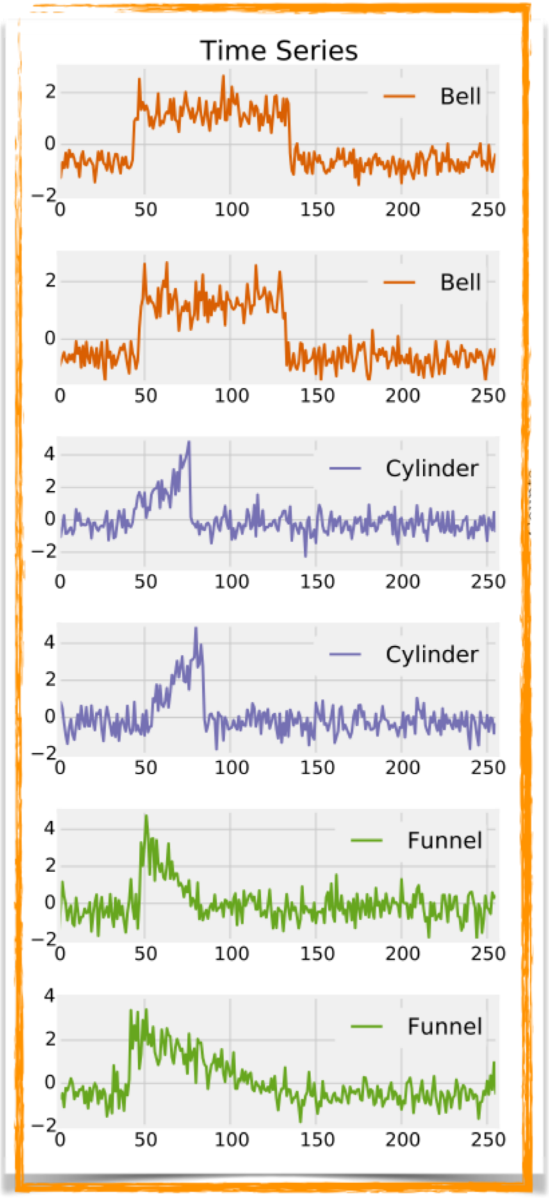
Land cover usage classification

- Typically, experts label some reference land cover samples apriori
- Then we can train classifiers in a supervised manner
- However, selecting an appropriate number of samples is crucial and very time consuming as it is typically done manually (and error-prone)
- Thus, acquiring reference data is very expensive and time consuming, thus large labelled datasets are rare and very valuable
- Luckily, we are given the labelled data and only need to worry about choosing the right classifier

Classification (I)

- Given a set D of samples and a set of classes C. A classifier is a function f: D→C
- Supervised learning:
    - Training data: Obtain a set S of samples with their classes
    - Feature Space: Function v mapping a sample to a vector of features.
    - Model the characteristics of the samples in each class
    - Encode the model in a classifier function f operating on the feature vector: v: D→V, and f: V→C
    - Classification: Compute f(v(d))

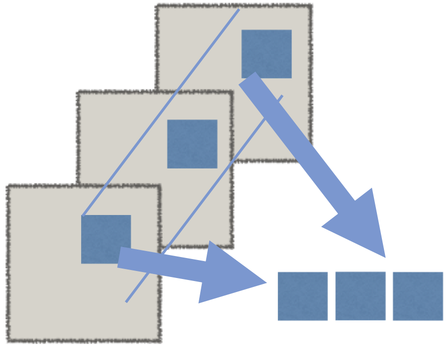
Dataset D



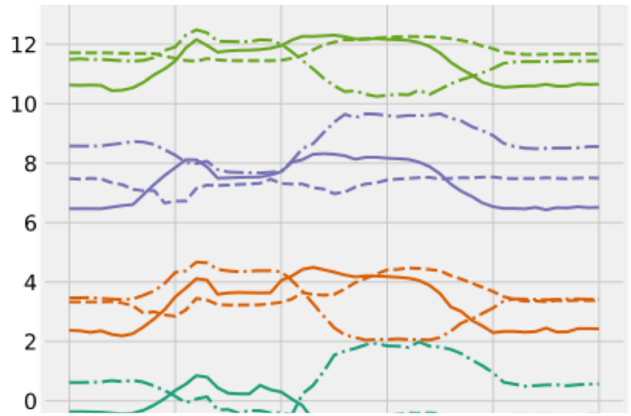
Classification (II)

- Non-time-series approaches do not capture model temporal dependencies and treat values independently

- Time series approaches explicitly capture temporal dependencies, which captures periodic changes over time (seasonality, agriculture, or harvesting)



- Multivariate ~: use all sensors as features but scalability might be a major issue (2d vector)
- Univariate ~: dimensionality reduction on features / use a single sensor or computed index (1d vector)



Time Series Classification TSC

- There are hundredths of base-classifiers
  - k-nearest neighbour, Naïve Bayes, Bayesian Networks, Graphical models, Decision Trees and Random Forests, Support Vector Machines, Neural Networks, …
- Differences when using different base-classifiers on the same data/representation are often astonishing small
- Including time series models into classification has a larger impact on accuracy than the choice of classifier
- Despite progress in technology and data availability, training time series models on large-scale data is very challenging in practice
- Overall, effectiveness of classification depends on many variables: labelled data, classifier, representation, feature selection and engineering, evaluation method

Time Series Approaches

- Time series approaches are composed of a time series representation and a classifier
- Representations of time series can be divided into:
  - Using the whole time series (global trends)
  - Using sub-sequences of a time series (local trends)
    - Shapelets: absence or presence of (seasonal) patterns
    - Bag-of-Patterns (Dictionaries): use frequency of occurrences

- Base-Classifiers can then trained on this (new) representation

Subsequence vs Whole Series

- Suppose, we wish to build a classifier to distinguish between two kinds of plants: what features should one use?
- The contour of a leaf can in fact be be interpreted as a time series
- Instead of comparing the entire shapes, it can be better to only compare small subsections ("Shapelets")
- Here: the defining difference is that Urtica dioica has a stem that connects to the leaf at almost 90 degrees

Ye, Lexiang, and Eamonn Keogh. "Time series shapelets: a new primitive for data mining." *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009.

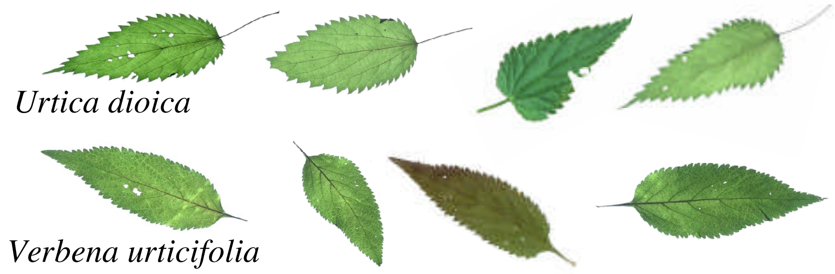


Figure 1: Samples of leaves from two species. Note that several leaves have the insect-bite damage

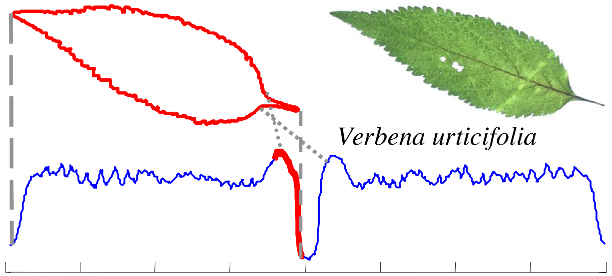
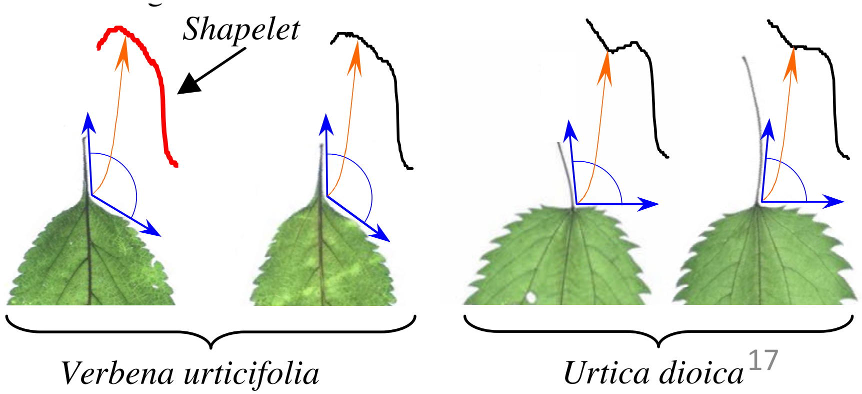
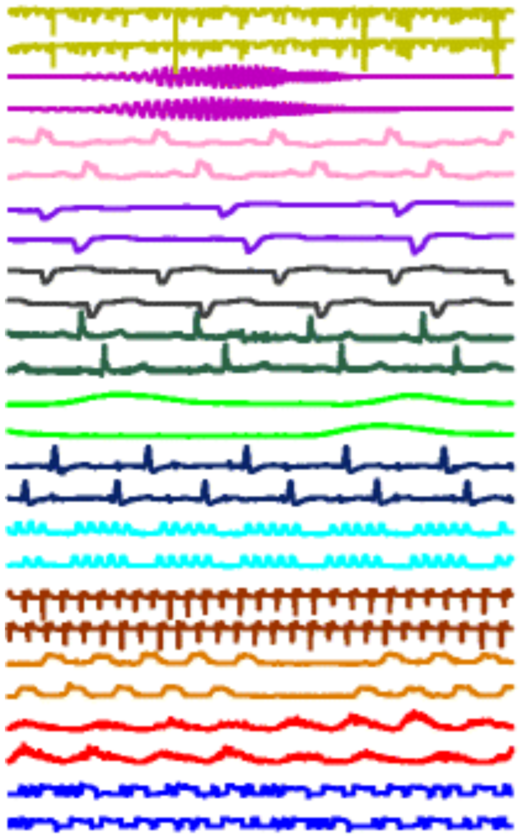


Figure 2: A shape can be converted into a one dimensional "time series" representation. The reason for the highlighted section of the time series will be made apparent shortly



Bag-of-Patterns:
single occurrence vs frequency of occurrences

- Many signals are inherently periodic/repetitive (heartbeats, network traffic, weather, …)
- We describe a signal by the frequency of occurrence of patterns
- Similar to the bag-of-words representation for documents, which is a histogram of word counts
- Problem: how to count the occurrences of real-valued-subsequences?

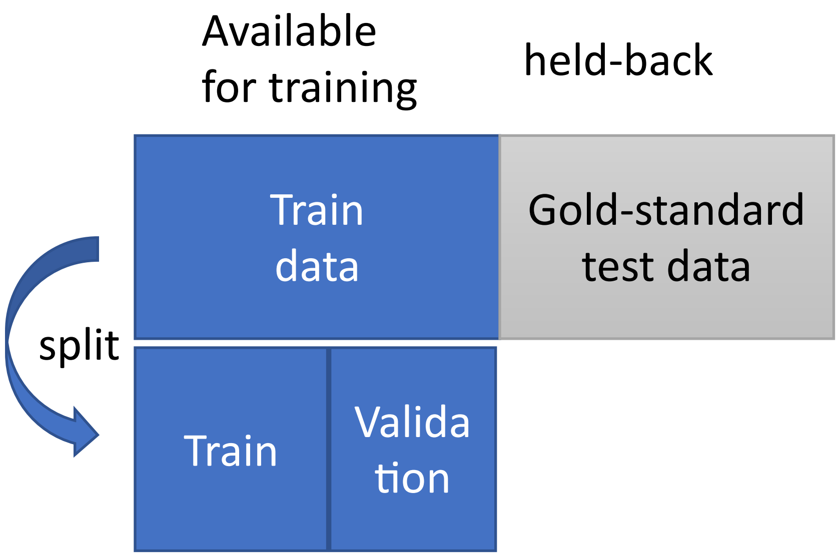


Lin, Jessica, Rohan Khade, and Yuan Li. "Rotation-invariant similarity in time series using bag-of-patterns representation." *Journal of Intelligent Information Systems* 39.2 (2012): 287-315.

What is a good Classifier?

- Problem: Finding a good classifier
  - Assigning as many samples as possible to their correct class
  - Involves proper feature engineering

- How do we know?
  - Use a (separate) gold standard test data set
  - Split the training data (beware of overfitting)
    - Train data for training the model
    - Validation data for evaluating the model



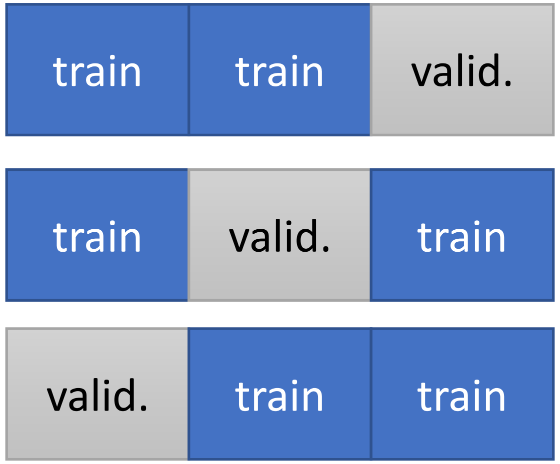
- A classifier f is the better, the more samples it assigns to their correct classes on the validation data

Beware of Overfitting

- Training data: Let S be a set of instances with their classes
  - We can easily build a perfect classifier for a train set S
  - Model: map a sample to its nearest neighbour in S using identity.
  - This will produce perfect results for every sample $s \in S$
  - What about samples $s \notin S$ (validation/test set)?
  - So, don't train a model and evaluate the model on the same data
- Overfitting
  - If the model strongly depends on S, f overfits – it will only work well if all future samples are very similar to the samples in S
  - You cannot detect overfitting when evaluation is performed on training set S only

Towards Overfitting

- f must generalize: Capture features that are typical for all samples in D, not only for the samples in S
- Still, often we only have a train set S for evaluation …
  - We need to extrapolate the quality of f to unknown samples
- Usual method: Cross-validation (leave-one-out)
  - Divide S into k disjoint partitions (typical: k=10)
    - Leave-one-out: k=|S|
  - Learn model on k-1 partitions and evaluate on the k-th
  - Performed k times, each time evaluating on another partition
  - Estimated quality on new samples = average performance over k runs



3-fold-cross-validation

What we’ll address here

- Working with massive pixel time series series
  - Feature engineering
  - (Time Series) Representations
  - (Time Series) Classification
  - Pre-processing
  - Scalability
  - Competition

Who should be here

- Master Informatik
  - Also: Wirtschaftsinformatik, Ms.Edu, Diplominformatik

- Ability to read English papers
- Ideally:
  - Knowledge on machine learning
  - Knowledge in statistics, probability theory, math
  - Or willingness to learn this

How it will work – competition part

- Every group (2-3 people) has to implement a (times series) classification method
  - Free choice pre-processing (normalization, interpolation, feature engineering, …)
  - Your topic: Every group uses (a) different classifier(s)
  - Published codes and libraries (Weka (Java), scikit-learn (Python), R, matlab,…) allowed
- I will release a training set in November
  - A set of pixels time series with landcover classes assigned
  - Program, build, test and optimize your model
  - Include approach description in seminar talk and presentation
- We will evaluate your method on held-back test data
  - You will be given an unlabeled test set
  - We will use an automated evaluation web platform (Kaggle)
  - Submissions will be possible in January
  - Small price for best average accuracy among all groups

Competition part

- A massive land cover pixel time series (TS) dataset
  - 46 geometrically and radio-metrically corrected images taken by FORMOSAT-2
  - Train data: 6.091.037 pixels TS, 2,4GB
  - Test data (hold-back): 2.614.122 pixels TS, 1,0GB
- In total 3x46 values per pixel time series
  - 46 time stamps between 06.02 and 29.11.2006
  - 3 surface reflectances: Near-Infra-Red, Red, Green
- Contains **missing values** ‚?‘
- Overall, 24 land cover classes, labelled by experts
- **Note: This data is provided for the class only and it has to be deleted once the seminar is over**

Competition: Excerpt of the data

- 0,226,57,57,285,47,50,401,49,49,408,43,51,459,45,53,320,20,25,460,69,82,?,?,?,?,?,?,338,119,109,?,?,?,331,135,102,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,269,95,81,266,100,81,283,90,76,283,86,71,282,84,71,298,80,70,?,?,?,313,80,67,311,79,70,356,86,76,332,83,74,?,?,?,185,91,66,199,87,66,174,85,59,175,88,62,?,?,?,156,60,47,?,?,?,131,47,40,159,41,32,186,27,32,140,21,31,167,23,24,188,41,33,182,43,30,116,30,26,135,28,23
- 0,156,51,52,154,30,36,292,30,33,316,30,43,417,38,44,288,16,19,433,64,70,?,?,?,?,?,?,325,111,95,?,?,?,308,109,79,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,248,64,60,246,75,60,255,68,55,251,59,49,294,64,53,250,59,56,?,?,?,283,64,52,281,67,53,309,64,55,288,65,57,?,?,?,130,54,35,126,50,43,121,50,34,121,56,43,?,?,?,76,32,24,?,?,?,82,28,28,65,26,23,89,18,20,69,15,19,74,17,16,108,32,27,77,15,16,69,15,13,91,18,18
- 0,254,71,63,285,66,64,347,58,57,365,54,57,378,58,59,248,30,30,431,69,82,?,?,?,?,?,?,337,112,105,?,?,?,308,104,89,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,279,70,64,271,65,63,288,69,65,287,66,62,282,63,60,290,64,63,?,?,?,303,66,59,305,64,58,347,65,65,330,65,63,?,?,?,235,77,65,229,73,61,202,76,59,239,79,61,?,?,?,234,55,49,?,?,?,239,67,62,267,54,51,277,47,53,300,46,54,295,41,43,296,58,59,266,60,54,210,58,52,175,46,44
- 0,221,64,56,302,48,54,397,44,48,412,41,48,450,43,53,312,20,25,467,72,83,?,?,?,?,?,?,325,128,114,?,?,?,315,155,110,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,277,73,69,271,79,73,290,73,65,285,71,62,294,71,62,305,70,65,?,?,?,317,71,65,320,72,64,366,74,68,338,65,64,?,?,?,187,85,64,193,82,59,166,81,58,165,84,62,?,?,?,162,61,49,?,?,?,138,49,41,189,47,37,207,33,33,165,25,31,229,27,25,204,42,39,199,40,27,124,28,23,132,28,26
- 0,153,51,50,157,28,34,324,32,37,354,30,44,426,36,46,312,19,19,433,64,69,?,?,?,?,?,?,323,122,104,?,?,?,289,125,87,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,262,55,58,257,64,57,268,62,55,261,56,52,280,61,49,256,57,56,?,?,?,276,63,51,296,64,56,326,60,57,300,62,57,?,?,?,136,59,41,138,58,42,132,56,39,127,60,48,?,?,?,95,37,27,?,?,?,85,28,32,80,30,24,75,18,17,77,14,19,86,16,16,108,31,29,79,18,18,71,14,21,88,18,18
- 0,262,69,63,325,59,68,336,55,57,369,50,57,375,51,57,262,27,27,403,67,78,?,?,?,?,?,?,320,110,104,?,?,?,303,117,89,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,255,78,70,266,75,66,266,69,65,278,71,63,248,67,62,285,71,64,?,?,?,300,71,61,286,69,63,337,66,66,303,65,63,?,?,?,203,74,64,215,74,63,196,74,59,193,84,64,?,?,?,212,60,48,?,?,?,230,74,63,262,59,55,297,53,61,290,46,50,298,42,51,275,53,53,273,53,56,199,47,42,179,41,37
- 0,235,61,57,323,48,54,397,47,49,418,41,49,454,37,53,315,20,26,462,69,85,?,?,?,?,?,?,323,128,113,?,?,?,317,149,109,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,?,284,72,68,282,74,67,292,72,65,300,67,61,297,68,61,307,67,64,?,?,?,329,66,62,333,69,62,383,67,66,351,65,63,?,?,?,185,87,64,181,82,59,163,74,56,159,85,61,?,?,?,138,61,47,?,?,?,144,51,41,186,47,41,203,36,34,165,29,31,221,27,28,204,40,38,199,37,26,118,26,20,132,27,19
- [...]

Direction: Features

- Scalability will be a major issue [1]!

Table 2
Variables derived from the spectrotemporal feature space by different approaches.

Variable type	Characteristics
Statistical metric	• Statistical summary of spectral values over one or more periods • Typically data informing seasonality or phenology • *Example*: average, maximum, minimum
Change metric	• Descriptive attribute of a temporal segment • Typically from annual data • *Example*: magnitude, duration, slope
Shape non-stationary	• Pattern of anniversary data • Can/cannot be characterized by parameters • Typically annual data
Shape stationary	• Periodic pattern of multiple values per year • Can be described by parameters • *Example*: sine or cosine based curve
Trend	• Non stationary • Annual or longer interval • Admits irregular intervals

Gómez, Cristina, Joanne C. White, and Michael A. Wulder. "Optical remotely sensed time series data for land cover classification: A review." ISPRS Journal of Photogrammetry and Remote Sensing 116 (2016): 55-72.

- Directions:
  - How to handle **missing values** (interpolation, …)?
  - Use additional indices: brightness, normalized difference vegetation index
  - Test additional features: minimum, maximum, average, range, …
  - Address Scalability: feature selection, dimensionality / noise reduction, sampling
  - Use a time series representation: Shapelets / Dictionary-based, …?

[1] Schäfer, Patrick, and Ulf Leser. "Benchmarking univariate time series classifiers." *Datenbanksysteme für Business, Technologie und Web (BTW 2017)* (2017).

A List of Possible Approaches

- **(non-time series) based-Classifiers**
  - SVM, logistic regression, random forests/decision trees, gradient boosting trees, XGBoost
- **Whole-Series-based Classifiers**
  - 1-NN Dynamic Time Warping
  - 1-NN Euclidean Distance
  - Proximity Forests
- **Shapelet-based Classifiers**
  - Univariate**:** Fast Shapelets (FS), Learning Shapelets (LS), Shapelet Transform (ST)
  - Multivariate: gRSF
- **Dictionary-based Classifiers**
  - Univariate**:** BoP, SAX VSM, TSBF, BOSS, BOSS VS, WEASEL
  - Multivariate**:** SMTS, WEASEL+MUSE, LPS
- **Deep Learning Classifiers**
  - ResNet**,** FCN, Encoder, MLP, Time-CNN, TWIESN, MCDCNN, MCNN, t-LeNet
- **Ensembles of Core Classifiers**
  - Univariate: EE PROP, COTE

A List of Non-Time Series Classifiers

Table 4
Strengths and weaknesses of algorithms used for large-area land cover characterization with time-series optical data.

Algorithm	Strengths/characteristics	Weaknesses
Artificial Neural Networks Non-parametric	• Manage well large feature space • Indicate strength of class membership • Generally high classification accuracy • Resistant to training data deficiencies—requires less training data than DT	• Needs parameters for network design • Tends to overfit data • Black box (rules are unknown) • Computationally intense • Slow training
Clustering *(partitioning)*	• Do not need previous knowledge • Do not need samples	• Cluster-class correspondence not assured • Complex identification of classes • Computationally intense
Decision trees Non-parametric	• No need of any kind of parameter • Easy to apply and interpret • Handle missing data • Handle data of different types (e.g. continuous, categorical) and scales • Handle non-linear relationships • Insensitive to noise	• Sensitive to noise • Tend to overfit • Not as good as others in large feature spaces • Large training sample needed
Gaussian Maximum likelihood Parametric	• Simple application • Easy to understand and interpret • Predicts class membership probability	• Parametric • Assumes normal distribution of data • Large training sample needed
Support Vector Machines Non-parametric	• Manages well large feature space • Insensitive to Hughes effect • Works well with small training dataset • Does not overfit	• Needs parameters: regularization and kernel • Poor performance with small feature space • Computationally intense • Designed as binary, although variations exist
Random Forests Non-parametric	• Capacity to determine variable importance • Robust to data reduction • Does not over-fit • Produces unbiased accuracy estimate • Higher accuracy than DT	• Decision rules unknown (black box) • Computationally intense • Needs input parameters (#trees and #variables per node)
Bagging	• Provides measures of classification confidence • Does not overfit	• Complex incomprehensible classifiers
Boosting	• Provides measures of classification confidence • Does not overfit • Robust to noise	• Stops if a classifier achieves zero training set error • Complex incomprehensible classifiers • Ineffective if excessive error in training sample

Gómez, Cristina, Joanne C. White, and Michael A. Wulder. "Optical remotely sensed time series data for land cover classification: A review." ISPRS Journal of Photogrammetry and Remote Sensing 116 (2016): 55-72.

General Literature

- On Landcover Classification - To be read by everyone
  - Gómez, Cristina, Joanne C. White, and Michael A. Wulder. "Optical remotely sensed time series data for land cover classification: A review." ISPRS Journal of Photogrammetry and Remote Sensing 116 (2016): 55-72.
  - Hostert, Patrick, et al. "Time series analyses in a new era of optical satellite data." Remote Sensing Time Series. Springer, Cham, 2015. 25-41.
- On Univariate Time Series Classification
  - Bagnall, Anthony, et al. "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances." Data Mining and Knowledge Discovery 31.3 (2017): 606-660.
- On Deep Learning for Time Series Classification
  - Fawaz, Hassan Ismail, et al. "Deep learning for time series classification: a review." arXiv preprint arXiv:1809.04356 (2018).
- Webseiten / Wettbewerbe:
  - TiSeLaC: Time Series Land Cover Classification Challenge
  - AALTD'16 Challenge : on multivariate time series data
  - http://timeseriesclassification.com : a website dedicated to univariate time series classifiers
- On classification / machine learning in general
  - Alpaydin: Introduction to Machine Learning, MIT press 2014
  - Ertel: Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung, Springer 2016

Allgemeine Hinweise

- Literaturrecherche ist notwendig
  - Die ausgegebenen Arbeiten sind Anker
  - Weiterführende Arbeiten müssen herangezogen werden
  - Auch Grundlagen nachlesen

- Kritisch lesen
  - Keine Angst vor nicht ganz zutreffenden Aussagen – solange gute Gründe vorhanden sind
  - Begründen und argumentieren
  - Kritikloses Abschreiben ist fehl am Platz

Next steps…

- Today: Introduction and group formation
- Friday 2.11.: Lecture “Primer on Time Series Analytics”
- Between today and Friday 2.11.: choose a topic
- Flash presentation (**7.12. 15-16 Uhr, RUD 25 4.410**):
  - Before 30.11.18: meet me to discuss topic
  - Present ideas and method in 5min
- Blockseminar (**1.2. 15-18 Uhr, RUD 25 4.410**)
  - Before 31.01.19: meet me to discuss slides
  - Present your topic (30-40min) at the Blockseminar
- Seminar Thesis before 31.03.2019!
  - write seminar thesis (~20 pages)

Topics and Groups

Topic	Assigned to (groups of 2-3)
(non-time series) based-Classifiers SVM, logistic regression, random forests/decision trees, gradient boosting trees, XGBoost, Bayesian methods	Alexej
Whole-Series-based Classifiers Dynamic Time Warping, Euclidean Distance, Proximity Forests	
Shapelet-based Classifiers Univariate: Fast Shapelets (FS), Learning Shapelets (LS), Shapelet Transform (ST) Multivariate: gRSF	
Dictionary-based Classifiers Univariate: BoP, SAX VSM, TSBF, BOSS, BOSS VS, WEASEL Multivariate: SMTS, WEASEL+MUSE, LPS	Arik
Deep Learning Classifiers ResNet, FCN, Encoder, MLP, Time-CNN, TWIESN, MCDCNN, MCNN, t-LeNet	
Ensembles of Core Classifiers Univariate: EE PROP, COTE	
...	

Siehe Webseite:
https://hu.berlin/landnutzung

Questions?

Hinweise zum Vortrag

- 30-40 Minuten plus Diskussion
- Klare Gliederung
- Themenauswahl: Lieber verständlich als komplett
- Bilder und Grafiken; Beispiele
- Eher Stichwörter als lange Sätze
- Vorträge können auch unterhaltend sein
    - Gimmicks, Rhythmuswechsel, Einbeziehen der Zuhörer, etc.
- Adressat sind alle Teilnehmer, nicht nur die Betreuer
- Technik: Laptop? Powerpoint? Apple?
- Siehe: https://hu.berlin/checkliste_seminar

Hinweise zur Ausarbeitung

- Eine gedruckte Version abgeben
  - Selbstständigkeitserklärung unterschreiben
- Eine elektronische Version schicken
- Referenzen: Alle verwendeten und nur die
  - Im Text referenzieren, Liste am Schluss
- Korrekt zitieren
  - Vorsicht vor Übernahme von kompletten Textpassagen; wenn, dann deutlich kennzeichnen
  - Aussagen mit Evidenz oder Verweis auf Literatur versehen
- Verwendung von gefundenen Arbeiten im Web
  - Möglich, aber VORSICHT
  - Eventuell Themenschwerpunkt verschieben – Betreuer fragen
- Siehe: https://hu.berlin/checkliste_seminar

Format

- Benutzung unserer Latex-Vorlage
- Nur eine Schriftart, wenig und konsistente Wechsel in Schriftgröße und –stärke
- Inhaltsverzeichnis
- Bilder: Nummerieren und darauf verweisen
- Referenzen:
  - [1] Yan, X., Yu, P. S. and Han, J. (2004). "Graph Indexing: A Frequent Structure-Based Approach". SIGMOD, Paris, France.
  - [YYH04] Yan, X., Yu, P. S. and Han, J. (2004). "Graph Indexing: A Frequent Structure-Based Approach". SIGMOD, Paris, France.
- Darf man Wikipedia zitieren?
  - Ja, aber nicht dauernd
- Siehe: https://hu.berlin/checkliste_seminar

If you want to know more…

Sehr geehrte Institutsmitglieder,

am Dienstag, den 30.10.2018, verteidigt ab 10.00 Uhr c.t. in Raum 4.410 Herr Arik Ermshaus seine Bachelorarbeit mit dem Titel

Deep Spatio-Temporal Time Series Land Cover Classification

Alle Interessierten sind herzlich eingeladen.

Beste Grüße,
Ulf Leser