Stance Detection: Concepts, Approaches, Resources, and Outstanding Issues

[Part 2]

ACM SIGIR 2021 Tutorial

DİLEK KÜÇÜK
Electrical Power Technologies Department
TÜBİTAK Energy Institute, Ankara-Turkey
dilek.kucuk@tubitak.gov.tr

FAZLI CAN
Computer Engineering Department
Bilkent University, Ankara-Turkey
canf@cs.bilkent.edu.tr

Part 2 Coverage: Outstanding Issues, Applications & Conclusions

1. **Outstanding Issues**
    1. Stance Detection in Data streams
    2. Context-sensitive stance detection
    3. Cross-lingual and multilingual stance detection
    4. Stance detection on non-textual data and robots

2. **Application Areas**
    1. Opinion surveys/polling
    2. Public health surveillance
    3. Information retrieval
    4. Stance summarization
    5. Rumour classification
    6. Fake news detection

3. **Concluding Remarks**

Outstanding Issues in Stance Detection

1. **Stance Detection in Data streams**
2. Context-sensitive stance detection
3. Cross-lingual and multilingual stance detection
4. Stance detection on non-textual data and robots

Stance Detection in Data Streams

- Longer part of this second portion of our tutorial presentation
  - Data Stream Properties
  - Static Stance Detection
  - Data stream stance detection
  - Prequential evaluation
  - Concept Drift
  - Online stance detection scenarios

What is Data Stream?

- A sequence of non-stopping temporal data
- 3 Vs:
  - Velocity
  - Volume
  - Variety
- Limitless data
- Limited memory
- Limited processing time
- Data distribution may change over time

Data Streams

- Social media posts
- News articles
- User comments
- Stock tickers
- Sensor data
- Intelligence reports
- Web clicks
- …

Search Result Personalization

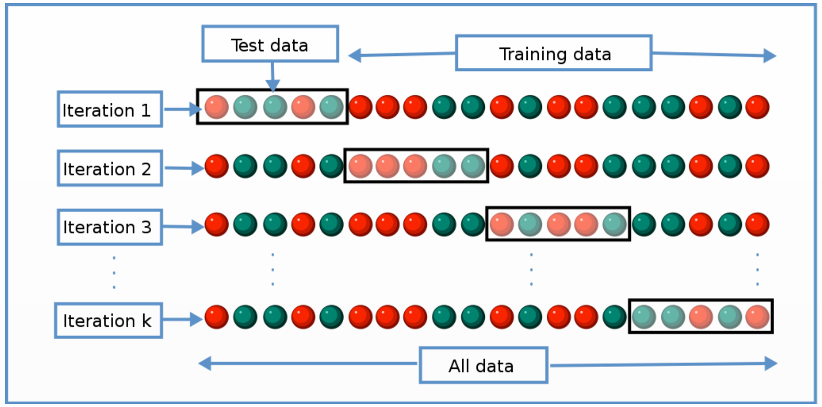
Eric Schmidt
Technical Advisor to Alphabet:

- "Google search will continue to become more personalized." (2012)
- Google combines several signals: Page rank, click, similarity
- Follows us to match our personal choices or influence our choices.
- The data stream we generate: The way we access pages, changes in our choices …

https://thenextweb.com/google/2012/02/28/eric-schmidt-google-search-will-continue-to-become-more-personalized/

Static Stance Detection

- Static model/hypothesis
- Training, Verification, and Testing
- K-folding: Cross validation
- Static hypothesis during an iteration
- Evaluation: Accuracy and other measures
- If used with temporal data: Training for old data items using future data items: may not change the results but wrong in principle.



https://en.wikipedia.org/wiki/Cross-validation_(statistics)

Static Data Stance Detection

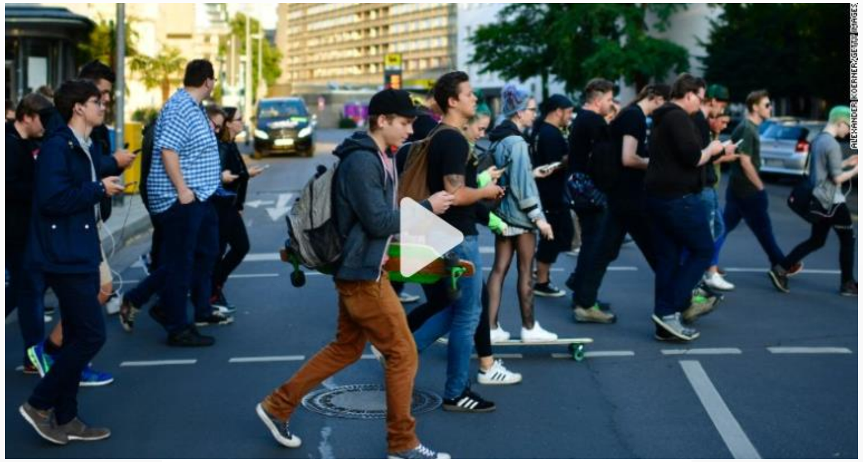
- Stance detection towards named entities
- Analysis of speeches (supporting bill/issue or against it)
- Support or opposition in debates

A. Hamdi et al. "Multilingual Dataset for Named Entity Recognition, EntityLinking and Stance Detection in Historical Newspapers," ***ACM SIGIR 2021 Proceedings,*** **2021.**

Sakala Venkata Krishna Rohit and Navjyoti Singh. 2018. Analysis of speeches in Indian parliamentary debates. arXiv preprint arXiv:1808.06834 (2018).

Data Stream Stance Detection

- People are Curious
- Dynamic results keep people up-to-date
  - People want to see what’s happening as soon as possible
  - Data stream analysis addresses human curiosity

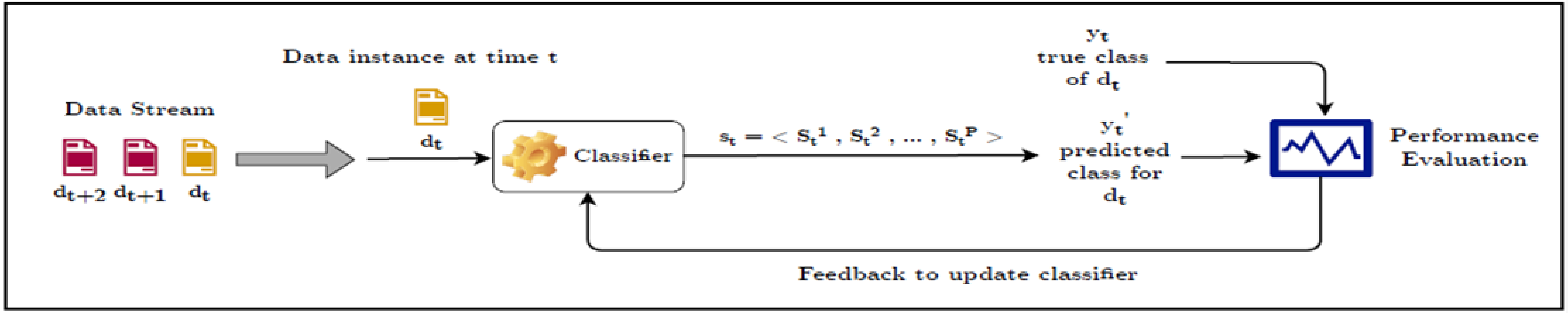


“Are we addicted to the Internet?”

https://edition.cnn.com/2017/07/29/us/smart-phones-crosswalks-hawaii-illegal/index.html

Online Data Stream Processing

- No separate train and test sets
- Use interleaved-test-then-train method: use each instance first to **test** the model, and **then** to **train** the model
- Continuous change of the classification model (variations possible)



Using Ensemble Methods

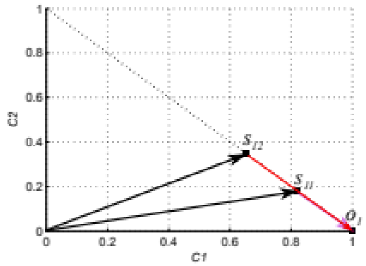
- Ensemble learning is a paradigm where multiple machine learning algorithm results are combined to get more accurate results.
- Each component can use a different model, different types of features (text, speech, video,…)
- Some Ensemble methods:
  1. Bagging: sampling with replacement and averaging
  2. Boosting: combine weak learners (better than random) to obtain a strong learner
  3. Stacking: a learner that learns classifiers: GOOWE
- Using these methods and voting systems like majority voting, the system can adapt to new changes.
- GOOWE: An example ensemble method.

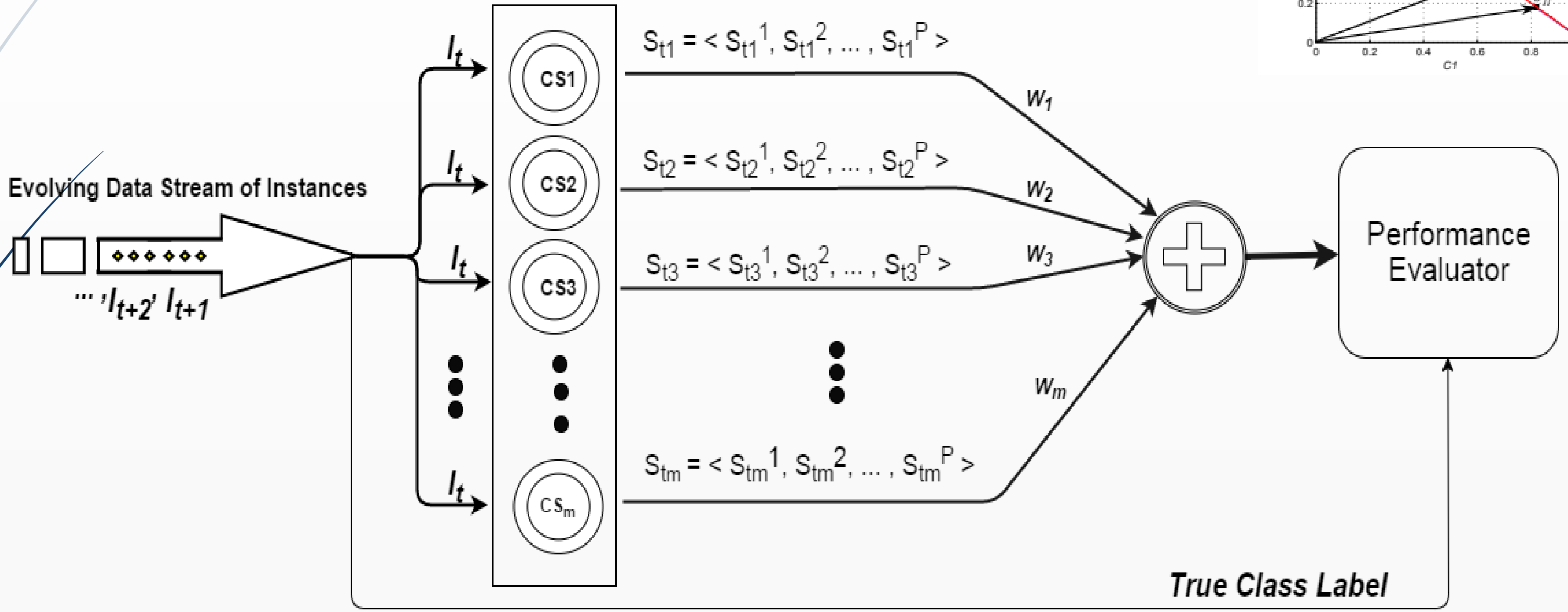
GOOWE: Geometrically Optimum ...

- Notations and Assumptions:
  - **p class labels** as C = (C1, C2… Cp)
    multi-class problem
  - **m classifier** systems as CS = (CS1, CS2… CSm)
  - For each of these Instances (I_i), every classifier system CS_j returns a set of scores as $S_{ij} = (S_{ij}^1, S_{ij}^2 \ldots S_{ij}^p)$

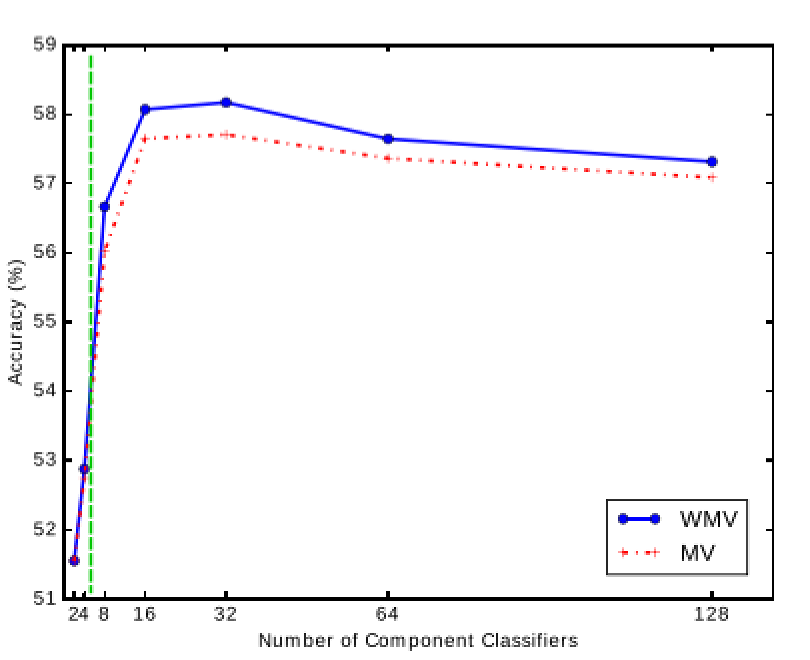
H. Bonab and F. Can, "GOOWE: Geometrically optimum and online-weighted ensemble classifier for evolving data streams", *ACM Transactions on Knowledge Discovery from Data*, 12(2), 1-25:33., 2018.

GOOWE: Geometrically Optimum …





Law of Diminishing Returns (m = p)



The highest effectiveness is observed much closer to the theoretically ideal m=p green line rather than the maximum number of components.

H. Bonab and F. Can, 2018.

Prequential Evaluation

- While calculating prequential accuracy, each data instance is used for two purpose: First testing then for training.
- Interleaved-test-then-train evaluation.
- Overall accuracy at time t, is the accumulated prequential accuracy calculated for all the stream data until time t.

V. M. A. Souza, D. F. Silva, J. Gama, G. E. A. P. A. Batista, "Data stream classification guided by clustering on nonstationary environments and extreme verification latency," *SDM*, pp. 873-881, 2015

Prequential Evaluation Assumption

- It assumes that the correct label is available right after testing.
- This assumption cannot be valid under all conditions.
- Some people argue that the above assumption is unrealistic.
- True label is not available: Extreme Verification Latency.

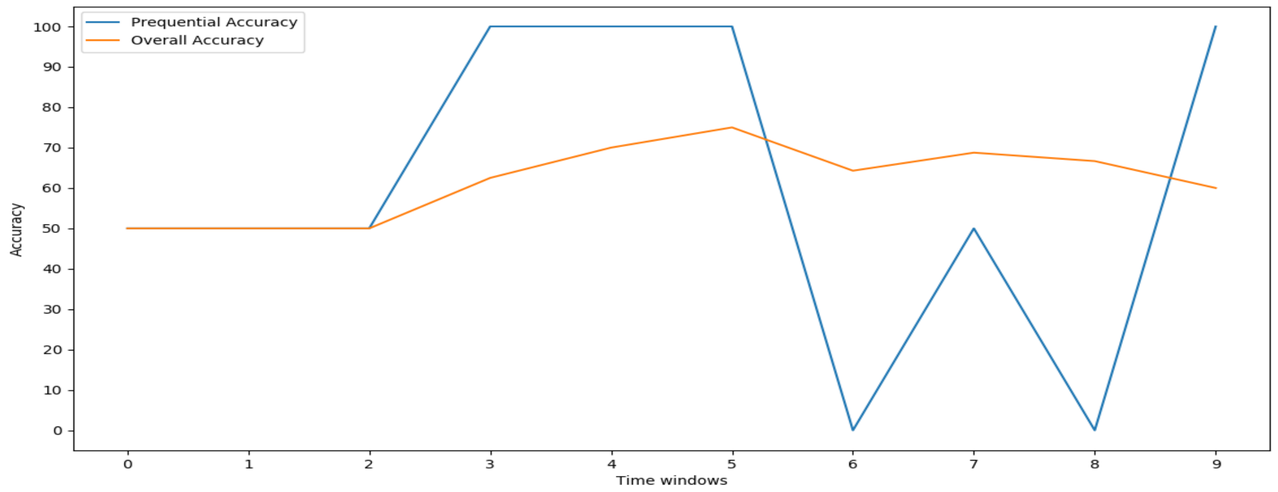
Vinícius M. A. de Souza, Diego Furtado Silva, João Gama, Gustavo E. A. P. A. Batista:
Data Stream Classification Guided by Clustering on Nonstationary Environments and Extreme Verification Latency. SDM 2015: 873-881

Prequential Evaluation: Example 1

There are 2 data instances in each time windows

Pediction	1,0	0,1	1,0	1,1	1,1	1,1	0,0	1,1	0,1	0,0

Prequential acc	50%	50%	50%	100%	100%	100%	0%	50%	0%	100%
Overall acc	50%	50%	50%	62.5%	70%	75%	64.28%	68.75%	66.66%	60%

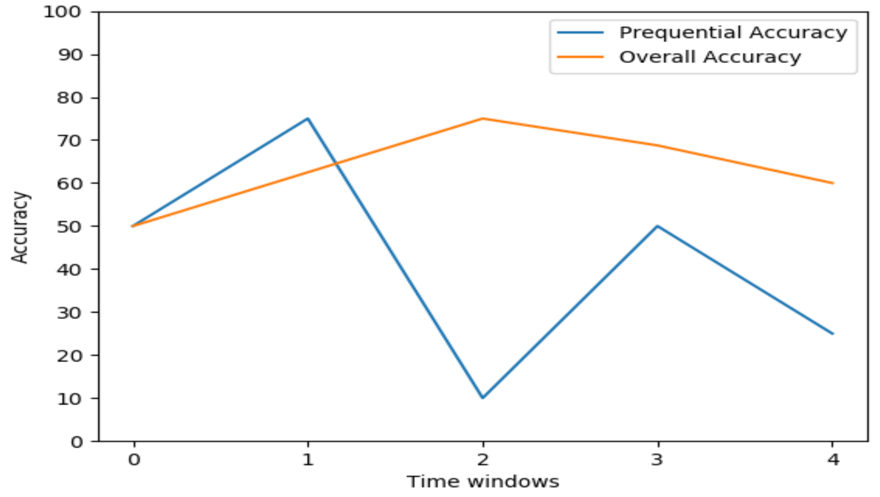


Prequential Evaluation: Example 2

There are 4 data instances in each time windows

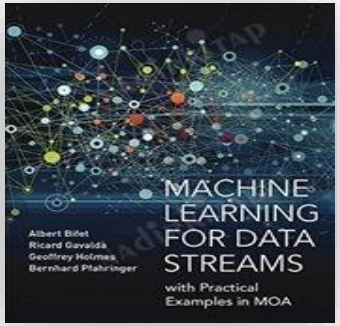
Pediction	1,0,0,1	1,0,1,1	1,1,1,1	0,0,1,1	0,1,0,0

Prequential acc	50%	75%	100%	50%	25%
Overall acc	50%	62.50%	75%	68.75%	60%



Analysis Tool: MOA - Massive Online Analysis

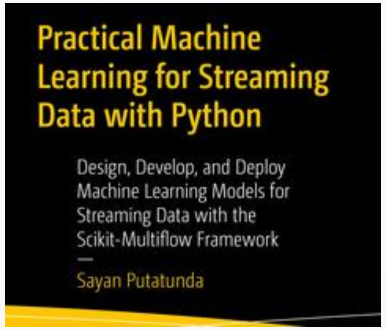
- MOA (Massive Online Analysis) is one of the popular frameworks for data stream analysis in JAVA.
- contains some of the most famous machine learning algorithms e.g. Hoeffding Tree, Naïve Bayes,…
- There are generators in MOA which simulate the stream environment by generating the data over time.
- https://moa.cms.waikato.ac.nz/



Analysis Tool: Scikit-multiflow

Scikit-learn is a famous machine learning framework in python

- However, scikit-learn doesn't contain any module for handling data streams
- Scikit-multiflow is a framework which is the equivalent of MOA in python
- It also has the capabilities of scikit-learn
- https://scikit-multiflow.github.io/



Dynamic Online Processing in Data Streams

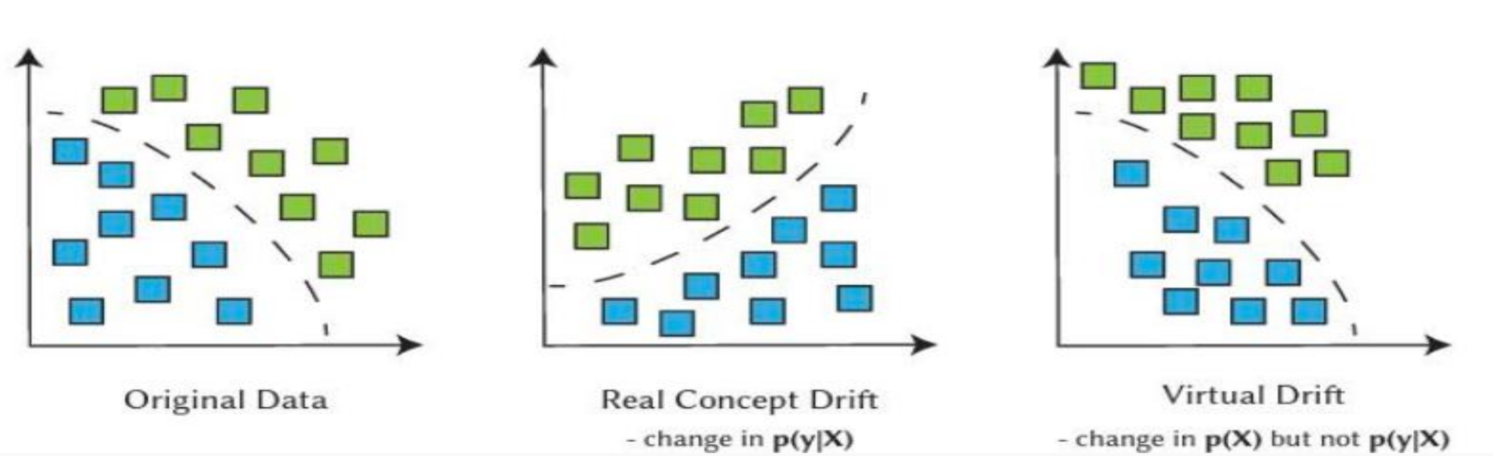
1. Process one example at a time and inspect it only once
2. Data is not stored in memory: Use a limited amount of memory
3. Be ready to predict at any time
4. Be able to react to **concept drift** in case of evolving data streams: Model should be able to catch the variations in data

Bifet, Albert, et al. MOA: Massive online analysis. *Journal of Machine Learning Research* 11. 1601-1604 (2010).

Concept Drift: What Happens?

- Y: Output, X: input, P(Y/X): class prediction for X
- Real Concept Drift
  - Classification boundary changes (X remains the same, P(Y/X) changes
- Virtual Concept
  - Data items of a class change: Virtual Concept Drift

Concept Drift: What Happens?



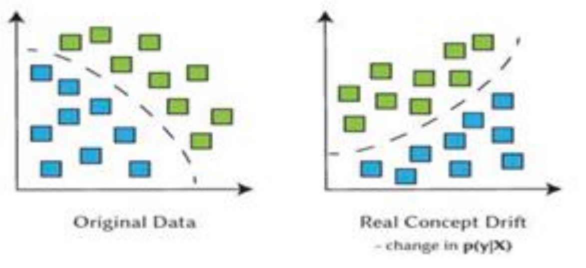
Decision boundary changes

Input characteristics changes

J. Gama, I. Žliobaite, A. Bifet, M. Pechenizkiy, and A. Bouchachia. 2014. A survey ˙ on concept drift adaptation. ACM Computing Surveys (CSUR) 46, 4 (2014), 44.

Real Concept Drift

- Decision boundary changes
  Change in p(y/X) with or without change in P(X)
  - Bachelor: Summer houses are interesting
  - Married: Dwelling houses become interesting



Virtual Concept Drift: What Happens?

- Input characteristic changes
- p(X) changes without affecting p(y|X)

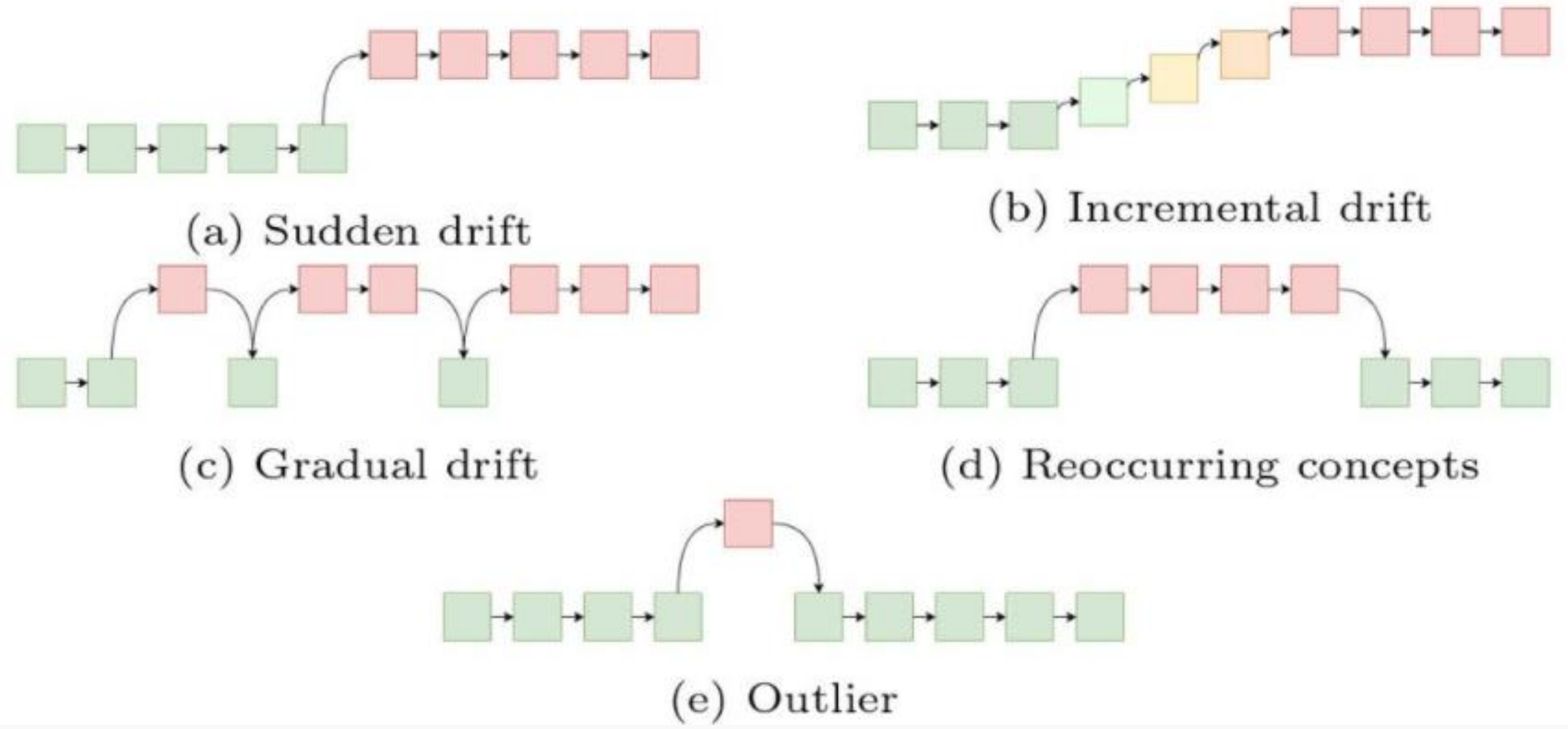
User favorite movie type is action
Action movies start to have comedy flavor (X changes) and user s[illegible] them (Y remains the same)

Drift Example

- “Example: Consider an online news stream of articles on real estate. The task for a given user is to classify the incoming news into relevant and not relevant. Suppose that the user is searching for a new apartment; then news on dwelling houses is relevant, whereas holiday homes are not relevant. If the editor of the news portal changes**, the writing style changes** as well, but the dwelling houses remain relevant for the user. This scenario **corresponds to virtual drift**. If, due to a crisis, more articles on dwelling houses come out and fewer articles on holiday homes do but the editor, the writing style, and the interests of the user remain the same, this situation corresponds to drift in prior probabilities of the classes. If, on the other hand, **the user has bought a house and starts looking for a holiday destination, dwelling houses become not relevant and holiday homes become relevant. This scenario corresponds to the real concept drift**. In this case, the writing style and the prior probabilities stay the same. It may happen that **all types of drifts may take place at the same time.**”

Gama, J., Zliobaite, I., Bifet, A., Pechenizkiy, M., Bouchachia, A. A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4), 44:1–44:37 (2014).

Concept Drift: How it Happens?



What to do for Concept Drift?

- Concept drift detection:
  - Prediction performance decrease
  - Using an Algorithm: supervised vs. unsupervised
- Concept drift handling
  - Active: Detect and retrain
  - Passive: Let the system handle

Lu, A. Liu, F. Dong, F. Gu, J. Gama and G. Zhang. Learning under Concept Drift: A Review. *IEEE Transactions on Knowledge and Data Engineering*, 31(12), 2346-2363 (2019).

Concept Drift Detection

- Concept drift detection algorithms find the drift point in a data stream.
- They usually detect the change based on the changes in the distribution of data over time.
- Detectors send an alarm in case of any changes and the system adapts to the changes based on a concept drift adaptation algorithm.

Concept Drift Handling

1. Update the model over time in certain time order. For example, update the model every week and learn the new data.
2. Incrementally learn the new data over time.
3. Assign more weights to incoming data right after the drift.
4. Restart learning from the drift point
5. Use ensemble of weak classifiers to adjust to the changes

Unsupervised Concept Drift Detection 1

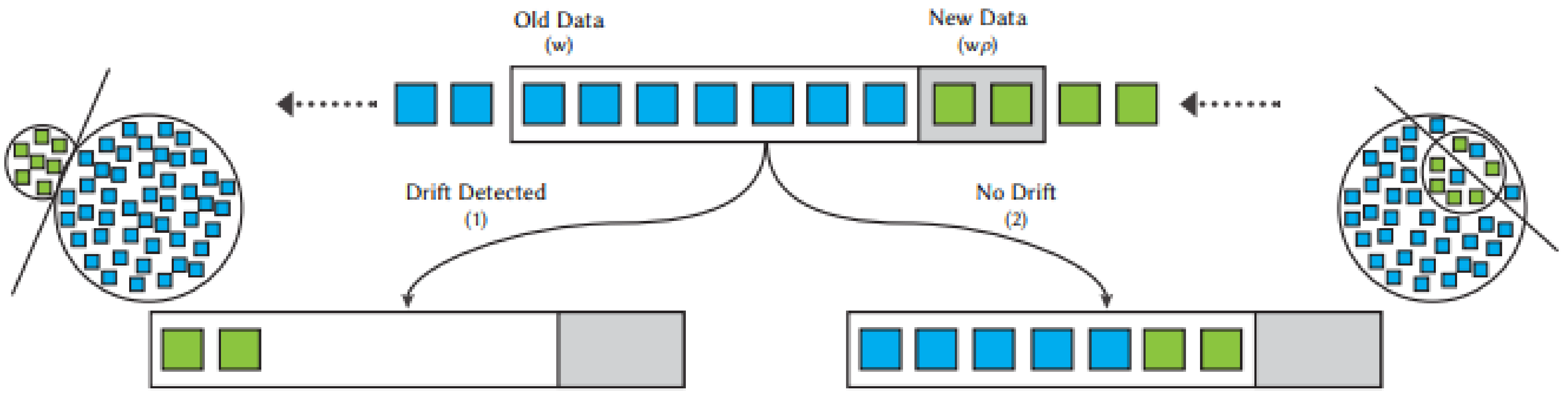
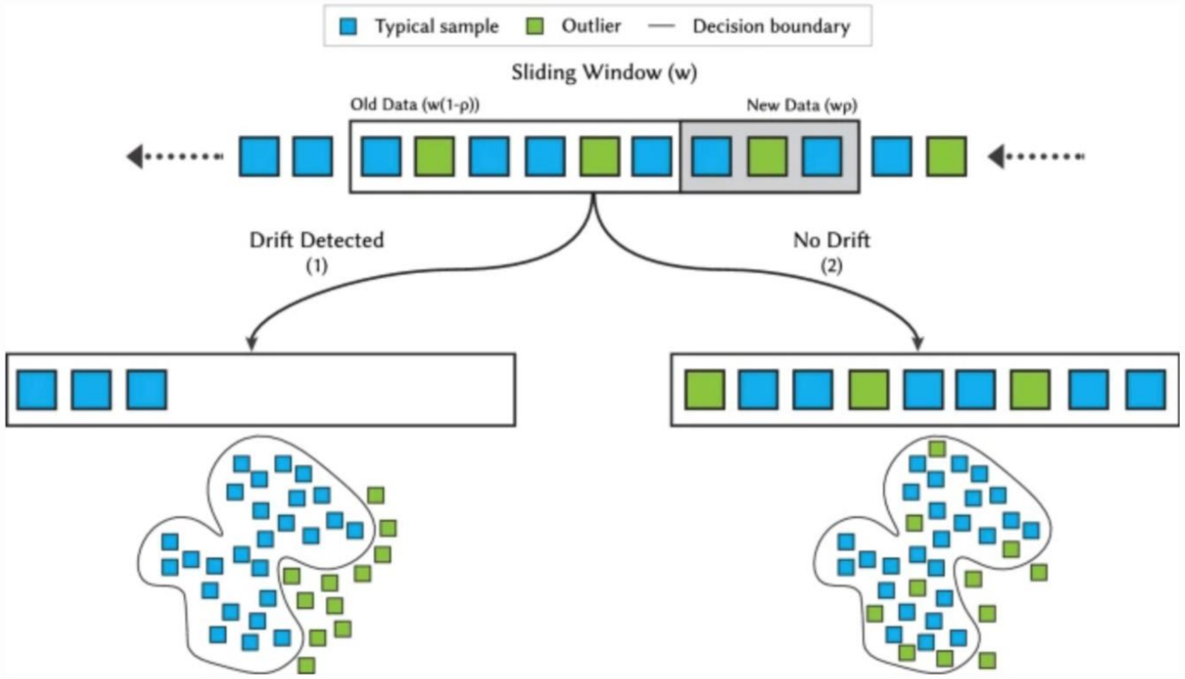


Figure 1: Drift detection workflow: (1): Drift detected. The old and the new data are separable. Samples from the old portion are discarded and partially filled with the samples from the new. (2): No drift. These sets are nested. The oldest $w\rho$ samples are removed and the window is shifted to left where the samples from the new fill the space that becomes empty.

Ö Gözüaçık, Baç üyükçakır, H. Bonab, F. Can, F. "Unsupervised concept drift detection with a discriminative classifier." (Short paper.) *The 28th ACM International Conference on Information and Knowledge Management*, 2019.

Unsupervised Concept Drift Detection 2



Drift detection workflow: (1): Drift detected. The percentage of outliers exceed the threshold (ρ). There is a change in the distribution of the data. Samples from the old portion are discarded and are partially filled with samples from the new data window. (2): No drift. There is no change in the data distribution. The oldest sample is removed and the window is shifted to the left, filling the empty space

Ö. Gözüaçık and F. Can, "Concept learning using one-class classifiers for implicit drift detection in evolving data streams." *Artificial Intelligence Review* 54 (5), 3725-3747.

Online Stance Detection Scenario

- No true labels are available: Extreme verification latency
- A set of initial data items with labels are available: Initial supervision
- An incremental concept drift environment: slow change in the input data items
- Use clustering and classification together

incremental change

Vinícius M. A. de Souza, Diego Furtado Silva, João Gama, Gustavo E. A. P. A. Batista:
Data Stream Classification Guided by Clustering on Nonstationary Environments and Extreme Verification Latency. SDM 2015: 873-881

Online Stance Detection Scenario (cont.)

1. Assign labels to a pool of data items by human annotation (c: number of labels: approve, disapprove, neutral)
2. Build an initial classification model using the initial labelled data
3. Divide the initial labelled data into c clusters (approve, disapprove, neutral)
4. Receive unlabelled new data items and insert them into a new pool
5. Detect the stance label of pooled data items with the available classification model
6. Use a clustering algorithm and obtain clusters for the pool
7. Map newly formed clusters to previously labelled clusters: By this way assign labels to new clusters (their members)
8. Replace the classification model using the labelled data items of the new clusters
9. Go to step 4

Vinícius M. A. de Souza, Diego Furtado Silva, João Gama, Gustavo E. A. P. A. Batista:
Data Stream Classification Guided by Clustering on Nonstationary Environments and Extreme Verification Latency. SDM 2015: 873-881

Use of Concept Drift Detection in This Scenario

If no concept drift is detected no need to renew the model.

Concept drift detection runs concurrently.

Advantage of Concept Drift Detection: Time efficiency: No delay in the results.

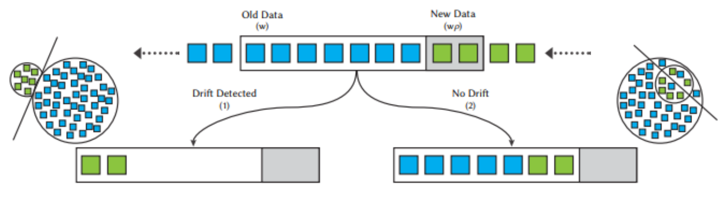
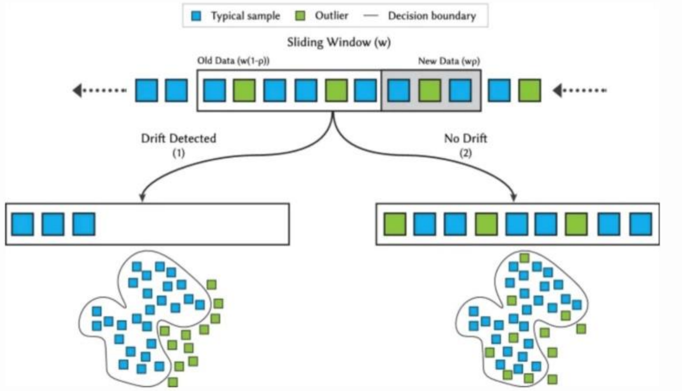


Figure 1: Drift detection workflow: (1): Drift detected. The old and the new data are separable. Samples from the old portion are discarded and partially filled with the samples from the new. (2): No drift. These sets are nested. The oldest $w\rho$ samples are removed and the window is shifted to left where the samples from the new fill the space that becomes empty.



Drift detection workflow: (1): Drift detected. The percentage of outliers exceed the threshold (ρ). There is a change in the distribution of the data. Samples from the old portion are discarded and are partially filled with samples from the new data window. (2): No drift. There is no change in the data distribution. The oldest sample is removed and the window is shifted to the left, filling the empty space

Vinícius M. A. de Souza, Diego Furtado Silva, João Gama, Gustavo E. A. P. A. Batista:
Data Stream Classification Guided by Clustering on Nonstationary Environments and Extreme Verification Latency. SDM 2015: 873-881

How to Incorporate Other Types of Concept Drifts

A small portion of labelled data is available spread in the data stream (1%, 5%, ... 100%).

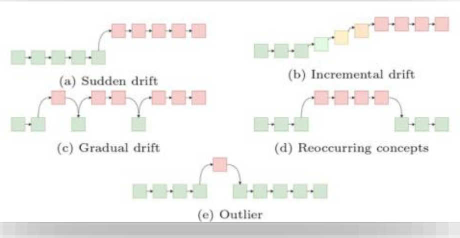
Availability of labelled data at different time instances makes system adaptable to different concept drifts.

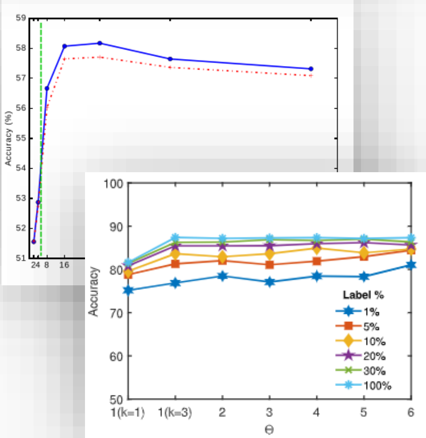
Clusters are constructed with maintenance.

Label propagation is used to label clusters.

Classification using clustering results.

An ensemble of k-NN classifiers are used.





S. U. Din, J. Shao, J. Kumar, W. Ali, J. Liu, J., and Y. Ye, "Online reliable semi-supervised learning on evolving data," *Information Sciences*, 525 (2020) pp. 153-171, 2020.

Fair Classifier Performance Evaluation

Evaluation Aspects: SeDiR

Scalability

- Applicable to data streams with different size & characteristics

Dynamism

- Adaptability to different types of concept drifts

Robustness

- Prequential evaluation: Missing labels, incorrect labels
- Missing/Incorrect features

Fair Classifier Performance Comparison

"… if it is possible to manipulate the experimental results using well-known statistical evaluation methods."

- Improper use of evaluation methodology (static dataset stance detection)
  - Problems in cross validation: How folding can change the results
  - Depending on the pool of classifiers results can change

Katarzyna et al. How to design the fair experimental classifier evaluation. Applied Soft Computing Journal, 104(2021),

Fair Classifier Performance Comparison (cont.)

Recommendations: not for data streams but still

- Use of accuracy and possible problems
- Use of different metrics and why
- Data set selection: how , what to pay attention
- Selection of proper statistical test

Katarzyna et al. How to design the fair experimental classifier evaluation. Applied Soft Computing Journal, 104(2021),

Evaluation of Online Stance Detection Applications

Different from lab conditions (in information retrieval remember Cranfield, TREC experimental environment)

User study: User satisfaction study

SDR: Scalability-Dynamism-Robustness

How to use it in classifier/application optimization: Depends on the user of the results

Katarzyna et al. How to design the fair experimental classifier evaluation. Applied Soft Computing Journal, 104(2021),

Outstanding Issues on Stance Detection

1. Stance Detection in Data streams
2. Context-sensitive stance detection
3. Cross-lingual and multilingual stance detection
4. Stance detection on non-textual data and robots

Context-sensitive Stance Detection

- How we see the world: Determined by our context.
- Context:
  - Google definition: “the circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood.”
  - Oxford Learner’s Dictionaries: “the situation in which something happens and that helps you to understand it”
- Principle of locality:
  - Physics: “an object is directly influenced only by its immediate surroundings.”
  - Computer Science: Temporal locality, Spatial locality.
- How we make decisions
  - How we proceed is affected by our locality/context.

Context-sensitive Stance Detection

- Using textual and contextual stance of tweets (Bharathi et al., 2020)
  - Contextual information
    - Network quote community
    - Network reply community
    - Network retweet community
    - Network friend community
    - User info bio
    - …
- Uses MLP (multilayer perceptron) and several models
- 5-fold cross validation: A problem that we may overlook.

B. Bharathi, J. Bhuvana, and N. N. A Balaji, "SSNCSE-NLP @ EVALITA2020: Textual and Contextual Stance Detection from Tweets Using Machine Learning Approach," CEUR-WS, Vol. 2765, Paper 151, 2020.

Cross-lingual and Multilingual Stance Detection

- Experiments need datasets.
- Datasets: Annotation ➔ difficult!
- How about borrowing from languages with datasets.
  - Annotated dataset in a given language (e.g., English) can be automatically translated into the target languages
  - Train a sentiment analysis model using recurrent neural networks with reviews in English. We then translate reviews in other languages and reuse this model to evaluate the sentiments (Can et al, 2018).

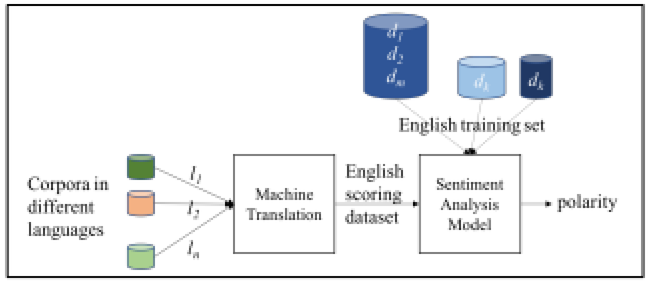


Figure 1: Multilingual sentiment analysis approach.

Can, E., F., Ezen-Can, A., and Can, F. "Multilingual sentiment analysis: An RNN-based framework for limited data," arXiv preprint arXiv:1806.04511 (2018).

Stance Detection on Non-textual Data and Robots

- Sentiment Analysis using facial, verbal, vocal input (Shrivastava et al, 2018) ➔ Stance detection!
- Joint stance detection from these different modalities

  - Presidential debate: Audience is give buttons for approve/disapprove. A facial sentiment/stance analysis?
  - Using such information by emotional robots (involves several ethical concerns)
  - Combining stance obtained from different modalities using an ensemble approach

V. Shrivastava, V. Richhariya, and V. Richhariya, "Puzzling Out Emotions: A Deep-Learning Approach to Multimodal Sentiment Analysis," in 2018 International Conference on Advanced Computation and Telecommunication (ICACAT), pp. 1-6, 2018.

Application Areas

1. Opinion surveys/polling
2. Public health surveillance
3. Information retrieval
4. Stance summarization
5. Rumour classification
6. Fake news detection

Opinion Surveys/Polling

- Topics include
  - Political/ideological/social debates
  - Product reviews
  - Elections/referendums

Traditional

Using social media, e.g. using Twitter (Sen et al., 2020)

You can never be sure!

- Emphasize the increasing importance of digital traces in polling
  - Performance across targets
  - Reliability
  - Direct vs. Indirect stance
- By means of automatic stance detection, whether a community is in favor of or against a topic of interest can be estimated, replacing (or complementing) the traditional practices of performing surveys/polls.

Sen, I., Flöck, F., Wagner, C. On the reliability of detecting approval of political actors in tweets. Proc. of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020, 1413-1426.

Public Health Surveillance

What is Public Health Surveillance?
"the continuous, systematic collection, analysis and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice. "WHO: Word Health Organization:

- The study by To, et al., (2021) uses stance detection for
  - Collection of information from social media that can be used for developing strategies to reduce anti-vaccination sentiments.
- The study by Zhang et al. (2017) uses stance detection for
  - Identification of controversial discussions regarding complementary and alternative medicine (CAM), stance detection on the posts in these discussions, and identification of the CAM therapies likely to trigger debates.

- **To, Q. G. Applying machine learning to identify anti-vaccination tweet during COVID-19 pandemic. International Journal of Environmental Research and Public Health, 2021, 18, 4069.**
- **Shaodian Zhang, Lin Qiu, Frank Chen, Weinan Zhang, Yong Yu, and Noémie Elhadad. 2017. We make choices we think are going to save us: Debate and stance identification for online breast cancer CAM discussions. In Proceedings of the International Conference on World Wide Web Companion. 1073–1081.**

Information Retrieval

1. Stance detection for search result personalization:
   - Determine user stance on different issues and provide results that would match their preferences.
   - It is possible to change user preferences.
   - Result: Eco room, Filter Bubble.
2. Answering multi-perspective queries:
   - “Is treatment x effective for disease y?” Answering such queries requires stances (support or oppose). (Sen et al., 2018).

Eli Pariser. 2011. The Filter Bubble: How the New Personalized Web is Changing What We Read and How We Think. The Penguin Press, New York.

Anirban Sen, Manjira Sinha, Sandya Mannarswamy, and Shourya Roy. 2018. Stance classification of multi-perspective consumer health information. In Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. 273–281.

Stance Summarization

- Observations:
  - Social media data items: continuous, endless, fast!
  - Hard to follow: Summarization of stances is needed (textual & visual)
- Studies:
  - Jang, Allan (2018): For tweets, defined as a ranking task, and use representative tweets as stance summary.
  - Wei et al. (2021): For movie critics and debates, first selects summary worthy documents, second stage uses MMR (maximal marginal relevance).

Myungha Jang and James Allan. 2018. Explaining controversy on social media via stance summarization. In Proceedings of the International ACM SIGIR Conference on Research & Development in Information Retrieval.

Penghui Wei, Jiaho Zhao, Wenji Mao, A Graph-to-Sequence Learning Framework for Summarizing Opinionated texts. IEEE/ACM Transactions on Speech and Language Processing Vol 29, 2021.

Stance Summarization

- Promising research area:
  - Google Scholar search: “Stance Summarization” *returns 16 results* (June 17, 2021).
  - Practical need.
  - Providing visual summaries and providing different views.
  - Named entity-based summarization.
  - Has a long history for text.

Myungha Jang and James Allan. 2018. Explaining controversy on social media via stance summarization. In Proceedings of the International ACM SIGIR Conference on Research & Development in Information Retrieval.

Penghui Wei, Jiaho Zhao, Wenji Mao, A Graph-to-Sequence Learning Framework for Summarizing Opinionated texts. IEEE/ACM Transactions on Speech and Language Processing Vol 29, 2021.

Rumour Classification

Rumour is defined as a piece of information that has not yet been verified.

They are misinformation with no intention of deceive. They can be spreaded for entertainment.

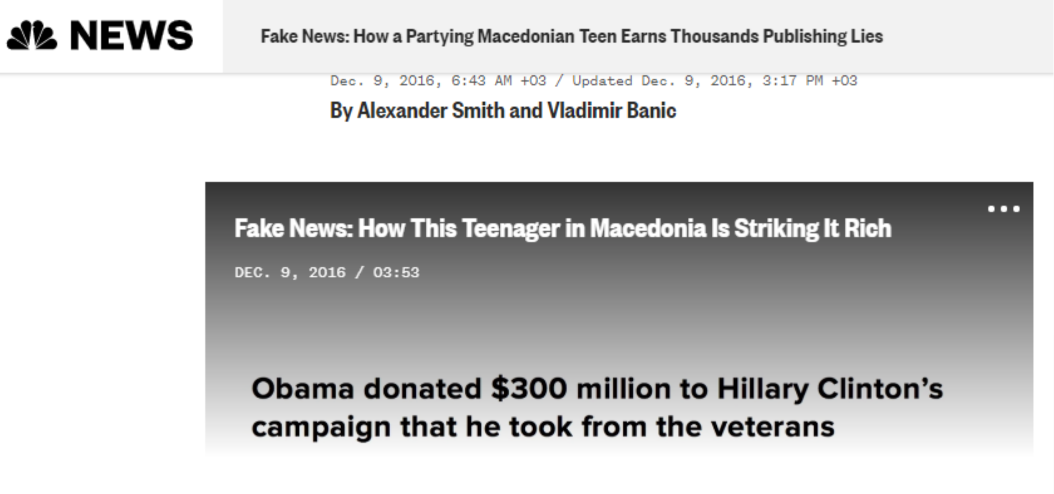
- The work by Zubiaga et al. (2018a) has four basic components in rumour processing
  - rumour identification,
  - tracking,
  - classification of rumour stance {Supporting, Denying, Querying, Commenting}
  - Veracity {True, False, Unverified}

Rumour stance detection and fake news stance detection can employ shared datasets.

Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. 2018a. Detection and resolution of rumours in social media: A survey. ACM Computing Surveys (CSUR) 51, 2 (2018), 32.

Fake News Detection

Growth of fake news? ➔ Effects on democracy and public trust.



Orson Welles, War of the Words, 1938

NBC News: https://www.nbcnews.com/news/world/fake-news-how-partying-macedonian-teen-earns-thousands-publishing-lies-n692451

Fake News Detection

They can be detected from their (Zhou, Zafarani, 2020)

- False knowledge they carry
- Writing style
- Propagation pattern
- Credibility of its resources

Immediate detection can be important (Ksieniewicz et al., 2020)

- Fake news detection in data streams

Zhou, X, Zafarani, R. A survey of fake news: Fundamental theories, detection methods, and opportunities, ACM Computing Surveys, 53(5): Article 109. September 2020.

Ksieniewicz, P. et al., Fake news detection from data streamsIJCNN, 2020.

Concluding Remarks

- Progress So Far
- Future Work
- Stance detection and social responsibility
- Resilient Systems

Progress So Far

- Several approaches
- Annotation Guidelines
- Datasets
- Evaluation Metrics
- Software and Tools
- Several Application Areas
- Several Future Research Possibilities

Future Work

- Common test collections: Larger and Different Languages
  - Replicable research
  - Objective comparison of results
- Synthetic data stream generation
  - Always in the wish list: Done in many classification tasks (see MOA)
  - Difficult: Natural language
  - Difficult: explicit vs. implicit stance, difficult for humans too

Future Work

- Libraries/Frameworks
  - Provides baselines
  - Open to additions in terms of algorithms and datasets
  - Making summarization a native part of them: seeing causality relationships
- Development of Online Systems
  - Online Stance Detection Evaluation
  - Definition & Evaluation Metrics: Are they the same as the lab results
  - How generalizable are the test results: theory to practice

Future Work

- From Individualized Implementations to More General Ones
  - Context-sensitive
  - Cross-lingual and multilingual
  - Non-textual data and robots
- All Data-Centric Approaches: Are they the only option?
  - Technology and amount of data are supportive
  - Is it like curve fitting for harvesting low-hanging fruits?
  - “There are so many low hanging fruits”
  - Neural networks used with causal modeling (J. Pearl, in Architects of Intelligence)

Ford, Martin, Architects of Intelligence: The Truth about AI from the People Building it, 2018, UK: Packt Publishing.

Stance Detection and Social Responsibility

- Computing technologies and ethics
  - “With great power comes great responsibility. Technology is now one of the most powerful forces shaping society, and we are responsible for it.” Moshe Vardi
  - Thinkers of the book *Architects of Intelligence*
  - What should we think about?
- Unintended) consequences of our research…
  - 2016 US elections
  - 2016 UK Brexit referendum
- What we **should not** create?”

Ford, Martin, Architects of Intelligence: The Truth about AI from the People Building it, 2018, UK: Packt Publishing. (Interviews with 23 scientists.)

Resilient Systems

- "Computing today is the 'operating system' of human civilization. We have the awesome responsibility."
- Systems must be resilient
  - Adaptable to disruptive environments
  - Robust to manipulation
  - Secure

Vardi MY (2020) Efficiency vs. resilience: what COVID-19 teaches computing. Commun ACM 63(5):9

References

1. A. Bifet, G. Holmes,R. Kirkby, and B. Pfahringer, "Massive online analysis," Journal of Machine Learning *Research* 11, May (2010), pp. 1601-1604.
2. B. Bharathi, J. Bhuvana, and N. N. A Balaji, "SSNCSE-NLP @ EVALITA2020: Textual and Contextual Stance Detection from Tweets Using Machine Learning Approach," *CEUR-WS*, Vol. 2765, Paper 151, 2020.
3. H. Bonab, and F. Can, "Less is more: A comprehensive framework for the number of components of ensemble classifiers," *IEEE Transactions on Neural Networks and Learning Systems*, 30(9), pp. 2735-2745, 2019.
4. H. Bonab, and F. Can, "GOOWE: Geometrically optimum and online-weighted ensemble classifier for evolving data streams," *ACM Transactions on Knowledge Discovery from Data*, 12(2), 25:1-25:33, 2018.
5. E. F. Can, A. Ezen-Can, A., and F. Can, "Multilingual sentiment analysis: An RNN-based framework for limited data," *arXiv preprint arXiv:1806.04511*, 2018.
6. S. U. Din, J. Shao, J. Kumar, W. Ali, J. Liu, J., and Y. Ye, "Online reliable semi-supervised learning on evolving data," *Information Sciences*, 525 (2020) pp. 153-171, 2020.
7. M. Ford, *Architects of Intelligence: The Truth about AI from the People Building it*, UK: Packt Publishing, 2018.
8. J. Gama, I. Zliobaite, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," *ACM Computing Surveys*, 46(4), 44:1–44:37, 2014.
9. Ö Gözüaçık, Baç üyükçakır, H. Bonab, and F. Can, F. "Unsupervised concept drift detection with a discriminative classifier." (Short paper.) *The 28th ACM International Conference on Information and Knowledge Management*, 2019.mber
10. Ö. Gözüaçık and F. Can, "Concept learning using one-class classifiers for implicit drift detection in evolving data streams." *Artificial Intelligence Review* 54 (5), 3725-3747.
11. Ahmed Hamdi et al. Multilingual Dataset for Named Entity Recognition, EntityLinking and Stance Detection in Historical Newspapers, ACM SIGIR 2021 Proceedings

References

12. Katarzyna et al. “How to design the fair experimental classifier evaluation,” *Applied Soft Computing Journal*, 104, 2021.
13. Ksieniewicz, P. et al., “Fake news detection from data streams,” IJCNN, 2020.
14. J. Lu, A. Liu, F. Dong, J. Gu, J. Gama, and G. Zhang, 2019, Learning under Concept Drift: A Review, *IEEE Transactions on Knowledge and Data Engineering*, 31(12), pp. 2346-2363, 2019.
15. MOA: https://moa.cms.waikato.ac.nz/
16. E. Pariser, *The Filter Bubble: How the New Personalized Web is Changing What We Read and How We Think*. The Penguin Press, New York, 2012.
17. Sakala Venkata Krishna Rohit and Navjyoti Singh. 2018. Analysis of speeches in Indian parliamentary debates. arXiv preprint arXiv:1808.06834 (2018).
18. Scikit multiflow: https://scikit-multiflow.github.io/
19. A. Sen, M. Sinha, S. S. Mannarswamy, and S Roy, “Stance classification of multi-perspective consumer health information,” *Proceedings of the ACM India Joint International Conference on Data Science and Management*, 2018.
20. I. Sen, F. Flöck, and C. Wagner, “On the reliability of detecting approval of political actors in tweets,” Proc. of the 2020 Conference on Empirical Methods in Natural Language Processing, pp. 1413-1426, 2020.
21. V. Shrivastava, V. Richhariya, and V. Richhariya, "Puzzling Out Emotions: A Deep-Learning Approach to Multimodal Sentiment Analysis," *Proceedings of International Conference on Advanced Computation and Telecommunication (ICACAT)*, pp. 1-6, 2018.
22. V. M. A. Souza, D. F. Silva, J. Gama, G. E. A. P. A. Batista, “Data stream classification guided by clustering on nonstationary environments and extreme verification latency,” *SDM*, pp. 873-881, 2015
23. G. I. Webb, R. Hyde, H. Cao, H.-L. Nguyen, and F. Petitjean,”Characterizing concept drift,” *CoRR abs/1511.03816, 2015.*
24. X. Zho and R. Zafarani, “A survey of fake news: Fundamental theories, detection methods, and opportunities,” *ACM Computing Surveys*, 53(5): Article 109. September 2020.

Thank You