PURDUE UNIVERSITY | Department of Computer Science

CS37300
Data Mining & Machine Learning

Ethics Issues in Data Mining

Profs. Steve Hanneke and Chris Clifton

28 April 2023

Indiana Center for Database Systems ™

PURDUE UNIVERSITY
Department of Computer Science

Ethics Issues for Data Mining & ML
What's the Problem?

- Privacy
  - Training data
  - Allowed uses
- Fairness
  - Inequitable outcomes
  - Variance in accuracy
- Data inaccuracy
- Explainability
- Redress
  - What if someone disputes results?

2

PURDUE UNIVERSITY
Department of Computer Science

Transparency

- Analyze and explain AI decision process
  - Very difficult
  - Likely only understandable to technology and domain experts
- Analyze and explain a decision
  - Input data analysis
  - Static explanation
  - Design/Code review and statistical analysis
  - Sensitivity analysis
  - Reverse-engineering the model

3

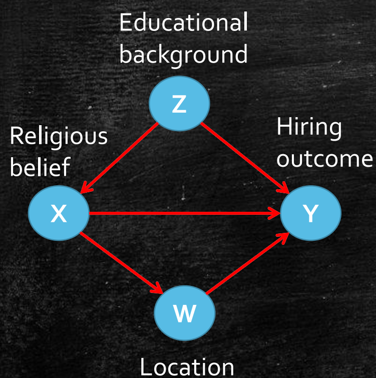
PURDUE UNIVERSITY
Department of Computer Science

GDPR Requirement: Transparency

- Article 13(2)(f), 4(2)(g): the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.
- Article 22(1) The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.
- Article 22(4) Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

4

Static Explanation through Causal Reasoning *(Junzhe Zhang and Elias Bareinboim AAAI'18)*



- The data analysis reveals that the total variation
$$E[Y|X=1]-E[Y|X=0] \ll 0$$

i.e., applicants of faith has lower chance of being hired.

- A frustrated applicant sues the company, claiming the disparity is due to:
  - Direct discrimination: the direct path $X \rightarrow Y$.
  - Indirect discrimination: the indirect path $X \rightarrow W \rightarrow Y$.
- The company argues the disparity is due to:
  - Difference in educational background: the spurious path $X \leftarrow Z \rightarrow Y$.

- Challenge: We do not have access to the code of the decision-making system (or the brains of the HR personnel in charge of hiring), so how to determine who is telling the truth?

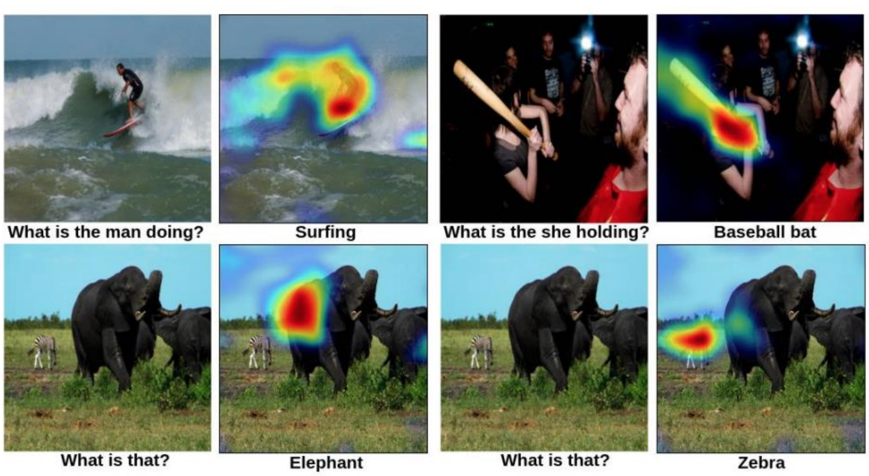
Fairness in Decision-Making, Zhang and Bareinboim, AAAI'18.

5

The University of Texas at Austin

WHAT STARTS HERE CHANGES THE WORLD

Visual Explanation



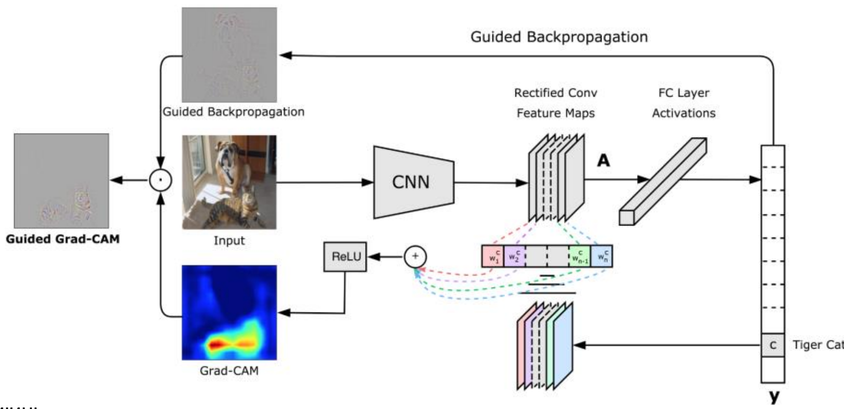
Dr. Nazneen Rajani

9

The University of Texas at Austin WHAT STARTS HERE CHANGES THE WORLD

Generating Visual Explanation

- *GradCAM* (Selvaraju et al., 2017) is used to generate heat-map explanations.



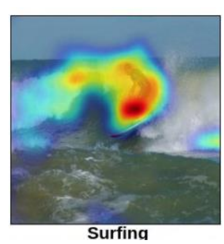
Dr. Nazneen Rajani

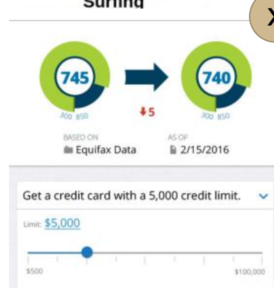
10

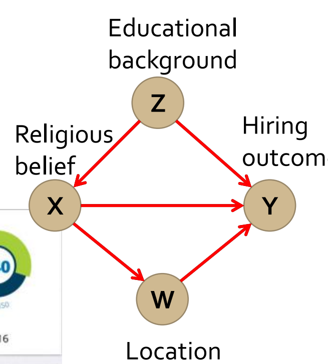
PURDUE UNIVERSITY Department of Computer Science

Are Explanations Accurate?

- Do these explanations really capture how decisions are made?
  - Sensitivity Analysis, Causal Reasoning
    - Explain outcome, not process
  - Heat maps
    - maybe?
- But does it matter?







11

Department of Computer Science

Emotional vs. Rational Decision-Making

- Humans have been shown to be emotional in their decision making
  - fMRI analysis of how decisions are made *(De Martino, Kumaran, Seymour, Dolan, Science 2006)*
- We rationalize our decisions
  - Explanations justify why we the decisions are good, not how we make them
- Is this good enough for explaining AI?
  - *Does this qualify as making ethical decisions?*

12

Department of Computer Science

Ethics Issues for Data Mining & ML *What's the Problem?*

- Privacy
  - Training data
  - Allowed uses
- Fairness
  - Inequitable outcomes
  - Variance in accuracy
- Data inaccuracy
- Explainability
- Redress
  - What if someone disputes results?

14

Top Ethical Issues
As presented at 2016 WEF

1. Unemployment
2. Distribution of machine-created wealth
3. Impact on human behavior/interaction
4. Guarding against mistakes
5. AI bias
6. Safety from adversaries
7. Protect against unintended consequences
8. How do we stay in control?
9. Robot rights

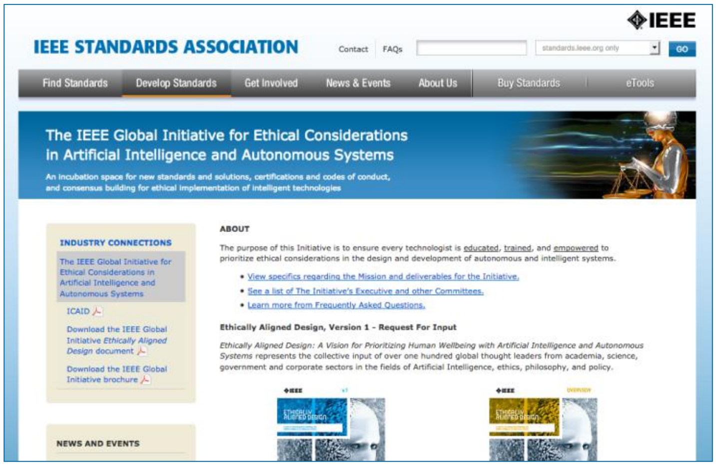
20

Ethical Issues:
AI Safety

- Multiple issues
  – Mistakes
  – Unintended consequences
  – Protection from adversaries
- Can we guarantee certain outcomes?
  – Rule out bad outcomes?

21

What do we do about it?
Standards and Best Practices

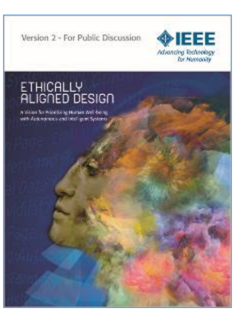


IEEE STANDARDS ASSOCIATION IEEE

25

Ethically Aligned Design

A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems



Version 2

- Launched December 2017 as a Request for Input
- Created by over 250 Global A/IS & Ethics professionals, in a bottom up, transparent, open and increasingly globally inclusive process
- Incorporates over 200 pages of feedback from public RFI and new Working Groups from China, Japan, Korea and more
- Thirteen Committees / Sections
- Contains **over one hundred twenty** key Issues and Candidate Recommendations

https://ethicsinaction.ieee.org/

IEEE STANDARDS ASSOCIATION IEEE

IEEE P70xx Standards Projects

IEEE P7000: Model Process for Addressing Ethical Concerns During System Design
IEEE P7001: Transparency of Autonomous Systems
IEEE P7002: Data Privacy Process
IEEE P7003: Algorithmic Bias Considerations
IEEE P7004: Child and Student Data Governance
IEEE P7005: Employer Data Governance
IEEE P7006: Personal Data AI Agent Working Group
IEEE P7007: Ontological Standard for Ethically Driven Robotics and Automation
IEEE P7008: Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems
IEEE P7009: Fail-Safe Design of Autonomous and Semi-Autonomous Systems
IEEE P7010: Wellbeing Metrics Standard for Ethical AI and Autonomous Systems
IEEE P7011: Process of Identifying and Rating the Trustworthiness of News Sources
IEEE P7012: Standard for Machines Readable Personal Privacy Terms

Related AI standards activities

- British Standards Institute (BSI) – BS 8611 *Ethics design and application of robots*
- **ISO/IEC JTC 1/SC 42 Artificial Intelligence**
  - ***SG 1 Computational approaches and characteristics of AI systems***
  - ***SG 2 Trustworthiness***
  - ***SG 3 Use cases and applications***
  - ***WG 1 Foundational standards***
- Jan 2018 China published "Artificial Intelligence Standardization White Paper."

PURDUE UNIVERSITY
Department of Computer Science

General Guidelines: FIPPs

Fair Information Practice Principles

- Transparency
  - Organizations should be transparent and notify individuals
- Individual Participation
  - Organizations should involve the individual in the process of using PII
- Purpose Specification
  - Organizations should specifically articulate the authority that permits the collection of PII
- Data Minimization
  - Organizations should only collect PII that is directly relevant and necessary
- Use Limitation
  - Organizations should use PII solely for the purpose(s) specified in the notice
- Data Quality and Integrity
  - Organizations should, to the extent practicable, ensure that PII is accurate, relevant, timely, and complete.
- Security
  - Organizations should protect PII (in all media) through appropriate security safeguards
- Accountability and Auditing
  - Organizations should be accountable for complying with these principles

NATIONAL STRATEGY FOR TRUSTED IDENTITIES IN CYBERSPACE - Appendix A

PURDUE UNIVERSITY
Department of Computer Science

Outline

- Use Cases
  - Autonomous weapons
  - Impact on people
- Limits of AI
  - Safety
- Decisions
  - Trolley problem
  - Discrimination
- Privacy
- Trust/Transparency
- Rights of AI
  - Legal personhood?
  - Intellectual Property?
- Ethical Reasoning
  - History

Rights of AI

- Can a machine have legal rights?
  – Animals do
  – Corporations, too
- What sort of rights should a machine have?
  – Rights of corporations?
  – Existence / not be “unplugged”?

33

Rights of AI: Intellectual Property

- U.S.: Only people own patents
  – Ever seen IBM or Google as the inventor in a patent?
  – Australia, South Africa have listed AI systems as inventors on patents
- What about copyright?
  – Australia, U.S. – copyright can only be awarded to a person
  – Is AI-generated art then public domain (uncopyrightable?)
    - Entertainment industry exploring this…

34

Rights of AI: Intellectual Property
UK Intellectual Property Office

"Consultation" updated 28 June 2022

- Copyright for AI-Generated Works
  - Currently protected under UK law
  - Plan no changes, but envisions potential for future changes
- Text/Data Mining
  - Plan to introduce copyright exception to allow TDM for any purpose
  - Still have safeguards for copyright holders
- Patent for AI Inventions
  - Currently AI *cannot* be held to be an inventor
    - But neither can human who invented the AI (unless involved in the invention)
  - As with copyright, no changes, but continue to review to support UK economic interests

35

Ethical Reasoning

- Ethical: Of or relating to moral principles
- Moral (of an action): having the property of being right or wrong, voluntary or deliberate and therefore open to ethical appraisal
- Ethical Reasoning in the context of AI (NSW Government):
  - A process of identifying ethical issues and weighing multiple perspectives to make informed decisions
  - Not about knowing right from wrong, but being able to think about and respond to a problem fairly, justly, and responsibly

36

Some suggestions

- Attend relevant talks
  - CS colloquium series (lists.purdue.edu – cs-colloq)
  - www.purdue.edu/critical-data-studies
- Data Ethics courses (a few)
  - ILS 23000: Data Science and Society: Ethical, Legal, Social Issues
  - PHIL 20700: Ethics for Technology, Engineering, and Design
  - PHIL 20800: Ethics of Data Science

37

CS37300
Data Mining & Machine Learning

Time Series Analysis

Profs. Steve Hanneke and Chris Clifton

28 April 2023

Mining Time-Series and Sequence Data

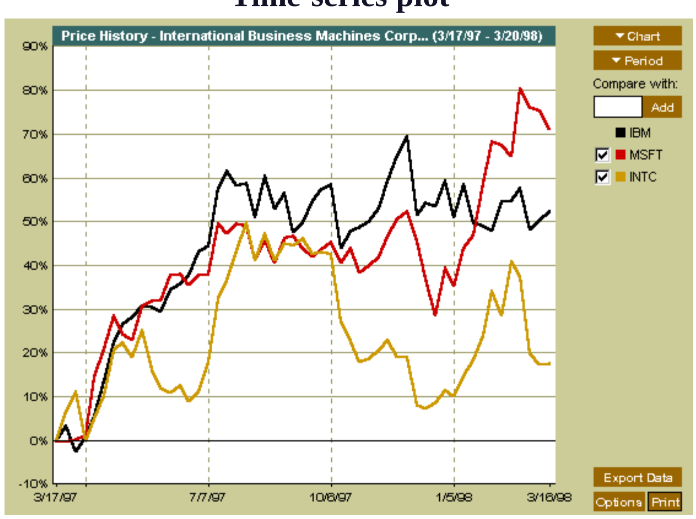
PURDUE UNIVERSITY
Department of Computer Science

- Time-series database
  - Consists of sequences of values or events changing with time
  - Data is recorded at regular intervals
  - Characteristic time-series components
    - Trend, cycle, seasonal, irregular
- Applications
  - Financial: stock price, inflation
  - Biomedical: blood pressure
  - Meteorological: precipitation

Mining Time-Series and Sequence Data

PURDUE UNIVERSITY
Department of Computer Science

Time-series plot



Mining Time-Series and Sequence Data: Trend analysis

- A time series can be illustrated as a time-series graph that describes a point moving with the passage of time
- Categories of Time-Series Movements
  - Long-term or trend movements (trend curve)
  - Cyclic movements or cycle variations, e.g., business cycles
  - Seasonal movements or seasonal variations
    - i.e., almost identical patterns that a time series appears to follow during corresponding months of successive years.
  - Irregular or random movements

41

Estimation of Trend Curve

- The freehand method
  - Fit the curve by looking at the graph
  - Costly and barely reliable for large-scaled data mining
- The least-square method
  - Find the curve minimizing the sum of the squares of the deviation of points on the curve from the corresponding data points
- The moving-average method
  - Eliminate cyclic, seasonal and irregular patterns
  - Loss of end data
  - Sensitive to outliers

42

Discovery of Trends in Time-Series (1)

PURDUE UNIVERSITY
Department of Computer Science

- Estimation of seasonal variations
  - Seasonal index
    - Set of numbers showing the relative values of a variable during the months of the year
    - E.g., if the sales during October, November, and December are 80%, 120%, and 140% of the average monthly sales for the whole year, respectively, then 80, 120, and 140 are seasonal index numbers for these months
  - Deseasonalized data
    - Data adjusted for seasonal variations
    - E.g., divide the original monthly data by the seasonal index numbers for the corresponding months

43

Discovery of Trends in Time-Series (2)

PURDUE UNIVERSITY
Department of Computer Science

- Estimation of cyclic variations
  - If (approximate) periodicity of cycles occurs, cyclic index can be constructed in much the same manner as seasonal indexes
- Estimation of irregular variations
  - By adjusting the data for trend, seasonal and cyclic variations
- With the systematic analysis of the trend, cyclic, seasonal, and irregular components, it is possible to make long- or short-term predictions with reasonable quality

44

Similarity Search in Time-Series Analysis

- Normal database query finds exact matches
- Similarity search finds data sequences that differ only slightly from the given query sequence
- Two categories of similarity queries
  - Whole matching: find a sequence that is similar to the query sequence
  - Subsequence matching: find all pairs of similar sequences
- Typical Applications
  - Financial market
  - Market basket data analysis
  - Scientific databases
  - Medical diagnosis

46

Data transformation

- Many techniques for signal analysis based on frequencies
  - Transform time series into frequency domain
- Usually data-independent transformations are used
  - The transformation matrix is determined a priori
    - E.g., discrete Fourier transform (DFT), discrete wavelet transform (DWT)
  - The distance between two signals in the time domain is the same as their Euclidean distance in the frequency domain
  - DFT does a good job of concentrating energy in the first few coefficients
  - If we keep only first a few coefficients in DFT, we can compute the lower bounds of the actual distance

47

PURDUE UNIVERSITY
Department of Computer Science

Multidimensional Indexing

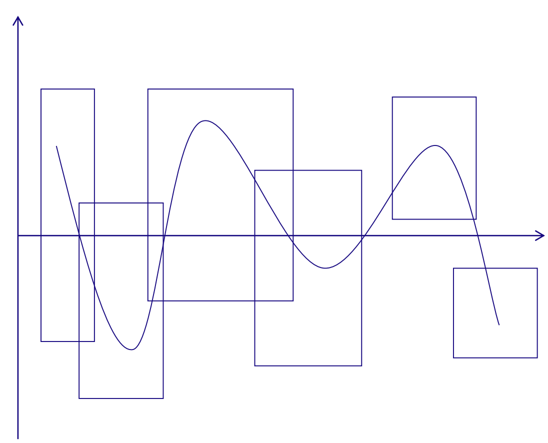
- Multidimensional index
  - Constructed for efficient accessing using the first few Fourier coefficients
- Use the index can to retrieve the sequences that are at most a certain small distance away from the query sequence
- Perform post-processing by computing the actual distance between sequences in the time domain and discard any false matches

48

PURDUE UNIVERSITY
Department of Computer Science

Subsequence Matching

- Break each sequence into a set of pieces of window with length *w*
- Extract the features of the subsequence inside the window
- Map each sequence to a “trail” in the feature space
- Divide the trail of each sequence into “subtrails” and represent each of them with minimum bounding rectangle
- Use a multipiece assembly algorithm to search for longer sequence matches



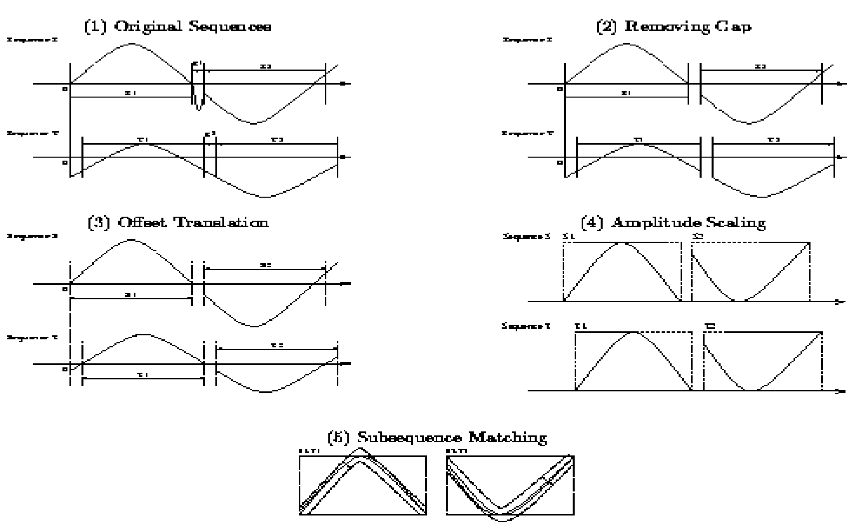
49

Enhanced similarity search methods

- Allow for gaps within a sequence or differences in offsets or amplitudes
- Normalize sequences with amplitude scaling and offset translation
- Two subsequences are considered similar if one lies within an envelope of ε width around the other, ignoring outliers
- Two sequences are said to be similar if they have enough non-overlapping time-ordered pairs of similar subsequences
- Parameters specified by a user or expert: sliding window size, width of an envelope for similarity, maximum gap, and matching fraction

PURDUE UNIVERSITY
Department of Computer Science

Similar time series analysis



Steps for Performing a Similarity Search

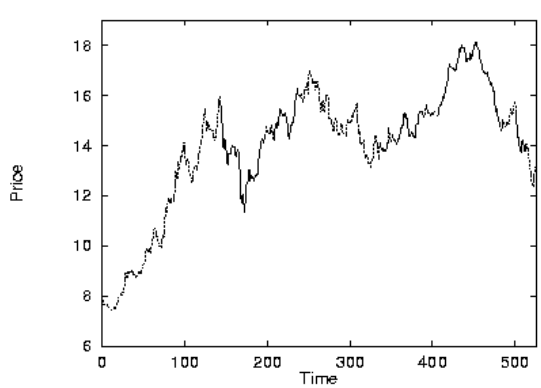
- Atomic matching
  - Find all pairs of gap-free windows of a small length that are similar
- Window stitching
  - Stitch similar windows to form pairs of large similar subsequences allowing gaps between atomic matches
- Subsequence Ordering
  - Linearly order the subsequence matches to determine whether enough similar pieces exist

52

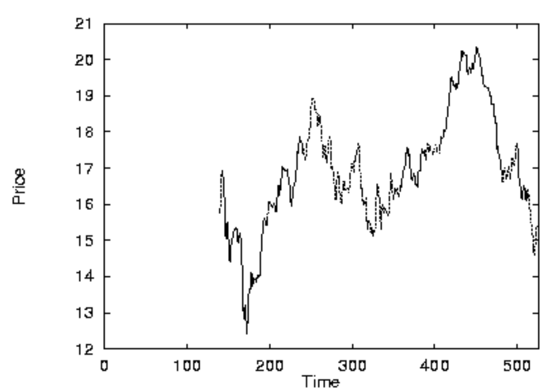
Similar time series analysis

Two similar mutual funds in different fund groups

VanEck International Fund



Fidelity Selective Precious Metal and Mineral Fund



53

Query Languages for Time Sequences

PURDUE UNIVERSITY
Department of Computer Science

- Time-sequence query language
  - Should be able to specify sophisticated queries like
    Find all of the sequences that are similar to some sequence in class A, but not similar to any sequence in class B
  - Should be able to support various kinds of queries: range queries, all-pair queries, and nearest neighbor queries
- Shape definition language
  - Allows users to define and query the overall shape of time sequences
  - Uses human readable series of sequence transitions or macros
  - Ignores the specific details
    - E.g., the pattern up, Up, UP can be used to describe increasing degrees of rising slopes
    - Macros: spike, valley, etc.

54

Sequential Pattern Mining

PURDUE UNIVERSITY
Department of Computer Science

- Mining of frequently occurring patterns related to time or other sequences
- Sequential pattern mining usually concentrate on symbolic patterns
- Examples
  - Renting “Star Wars”, then “Empire Strikes Back”, then “Return of the Jedi” in that order
  - Collection of ordered events within an interval
- Applications
  - Targeted marketing
  - Customer retention
  - Weather prediction

56

Mining Sequences (cont.)

Customer-sequence

CustId	Video sequence
1	{(C), (H)}
2	{(AB), (C), (DFG)}
3	{(CEG)}
4	{(C), (DG), (H)}
5	{(H)}

Map Large Itemsets

Large Itemsets	MappedID
(C)	1
(D)	2
(G)	3
(DG)	4
(H)	5

Sequential patterns with support > 0.25
{(C), (H)}
{(C), (DG)}

57

Sequential pattern mining: Cases and Parameters

- Duration of a time sequence *T*
  - Sequential pattern mining can then be confined to the data within a specified duration
  - Ex. Subsequence corresponding to the year of 1999
  - Ex. Partitioned sequences, such as every year, or every week after stock crashes, or every two weeks before and after a volcano eruption
- Event folding window *w*
  - If $w = T$, time-insensitive frequent patterns are found
  - If $w = 0$ (no event sequence folding), sequential patterns are found where each event occurs at a distinct time instant
  - If $0 < w < T$, sequences occurring within the same period *w* are folded in the analysis

58

Sequential pattern mining: Cases and Parameters (2)

- Time interval, int, between events in the discovered pattern
  - *int* = 0: no interval gap is allowed, i.e., only strictly consecutive sequences are found
    - Ex. "Find frequent patterns occurring in consecutive weeks"
  - *min_int* ≤ *int* ≤ *max_int*: find patterns that are separated by at least *min_int* but at most *max_int*
    - Ex. "If a person rents movie A, it is likely she will rent movie B within 30 days" (*int* ≤ 30)
  - *int* = *c* ≠ 0: find patterns carrying an exact interval
    - Ex. "Every time when Dow Jones drops more than 5%, what will happen exactly two days later?" (*int* = 2)

59

Episodes and Sequential Pattern Mining Methods

- Other methods for specifying the kinds of patterns
  - Serial episodes: A → B
  - Parallel episodes: A & B
  - Regular expressions: (A | B)C*(D → E)
- Methods for sequential pattern mining
  - Variations of Apriori-like algorithms, e.g., GSP
  - Database projection-based pattern growth
    - Similar to the frequent pattern growth without candidate generation

60

PURDUE UNIVERSITY
Department of Computer Science

Periodicity Analysis

- Periodicity is everywhere: tides, seasons, daily power consumption, etc.
- Full periodicity
  - Every point in time contributes (precisely or approximately) to the periodicity
- Partial periodicity: A more general notion
  - Only some segments contribute to the periodicity
    - Jim reads NY Times 7:00-7:30 am every week day
- Cyclic association rules
  - Associations which form cycles
- Methods
  - Full periodicity: FFT, other statistical analysis methods
  - Partial and cyclic periodicity: Variations of Apriori-like mining methods

61