

# Ausarbeitung zum Proseminar

Jan Grau

Martin-Luther-Universität Halle–Wittenberg

11. Juli 2007

# Outline

- 1 Vorbereitungen
- 2 Gliederung
- 3 Inhalte
- 4 Formulierung
- 5 Standing on the shoulders of giants

# Outline

- 1 Vorbereitungen
- 2 Gliederung
- 3 Inhalte
- 4 Formulierung
- 5 Standing on the shoulders of giants

# Literatur (Einstieg)

- Zuerst Überblick über Themenbereich schaffen
  - Lehrbücher (Bibliothek,...)
  - Überblicksartikel (Reviews)
  - auch: Wikipedia
  - Ziel:
    - in das Thema "reinkommen"
    - Überblick über/Referenzen auf detaillierte/speziellere Literatur
    - nicht: zitierfähige Inhalte

# Literatur (Inhalte)

- Dann Inhalte konkretisierten
  - Lesen, lesen, lesen
  - wenn möglich: verschiedene Sichtweisen
  - evtl. Excerpte
  - möglichst kurze Zusammenfassung (5 Sätze) der wichtigsten Inhalte jedes Papers
    - ⇒ hilft später beim Finden der Referenzen
  - evtl. Literatur priorisieren

# Technisches

## Vor dem Schreiben

- Welches Schreib-/Satzprogramm?  
Word, OpenOffice, InDesign, L<sup>A</sup>T<sub>E</sub>X, ...
  - Viele oder wenige Formeln
  - Viel oder wenig Literatur
  - Eigene Erfahrung
  - Layout, Seitenspiegel, Druck
- Literaturverwaltung
  - EndNote
  - verschiedene BibTeX-Frontends (BibDesk, kBib, ...)
  - Lohnt sich bei größerer Zahl Referenzen

# Outline

- 1 Vorbereitungen
- 2 Gliederung**
- 3 Inhalte
- 4 Formulierung
- 5 Standing on the shoulders of giants

# Gliederung

- Standardgliederung:
  - (Abstract)
  - Einleitung
  - Methoden
  - Ergebnisse
  - Diskussion
  - Zusammenfassung/Conclusion
  - Referenzen
- begründete Abweichungen
- auch: Ergebnisse und Diskussion gemeinsam
- hier: Literatarbeit  $\Rightarrow$  Ergebnisse?  
Aber: Beispiele, Anwendungen, Ergebnisse anderer, ...

# Abstract

- bei wissenschaftlichen Artikeln
- (meist sehr) kurze Zusammenfassung der gesamten Arbeit
- wichtigste Ziele, Methoden und Ergebnisse nennen
- meist keine Erklärungen, Definitionen, ...

# Einleitung

- “Aufhänger”
- Anwendungen
- Relevanz des Themas darlegen
- “Standortbestimmung”: Was war bisher, was ist aktuell, was ist neu? (weniger bei Literatuarbeit)
- evtl. grundsätzliche Konzepte
- evtl. Abgrenzung von anderen Themen
- evtl. kurzer Abriss der Inhalte

# Methoden

- bei uns meist: Mathematik
- Voraussetzungen formulieren, Annahmen nennen
- Ziel der Methode nennen
- Gleichungen liefern und erklären (!)
- Pseudo-code liefern und erklären (!)
- Abbildungen als Hilfestellung
- auch: Einordnen der Methode
  - Bsp. k-Means:
    - unüberwacht, Clustern, harte Zuordnung
    - nicht: überwacht, Klassifikation, weiche Zuordnung

... für die Ausarbeitung lieber mehrere Teile mit den einzelnen Methoden, da Ergebnisse und Diskussion (wahrscheinlich) nicht vorhanden

# Ergebnisse

- falls vorhanden
- Ergebnisse darstellen durch
  - Tabellen
  - Abbildungen, Plots
  - Text
- Abbildungen und Tabellen erklären (!)
- noch keine Interpretation

# Diskussion

- falls Ergebnisse vorhanden
- Interpretation
- Begründungen
- auf Methodik Bezug nehmen

Alternative: gemeinsam mit Ergebnissen

- Darstellung des Ergebnisses und Diskussion wecheln sich ab
- oft günstiger um “Argumentationsstrang” aufzubauen

# Zusammenfassung/Conclusion

- **gesamte** Arbeit zusammenfassen
- keine Angst vor Wiederholdungen!
- Teile aus der Einleitung wieder aufgreifen  
Viele Menschen lesen zunächst Einleitung und Zusammenfassung
- wichtigste Ergebnisse wiederholen
- auf Vorteile/Nachteile der Methodik Bezug nehmen
- evtl. Schlussfolgerungen für weitere Arbeit
- evtl. Ausblick
- keine “eigene Meinung”

# Referenzen

... dazu später mehr

# Beispielgliederung (Sequenzdaten)

- Einleitung
- Darstellung von Sequenzen
  - Konsensussequenzen
  - Sequenzlogos
- Position weight matrix
  - Modellierung und Definition
  - Schätzer
- Weight array matrix
  - Modellierung und Definition
  - Schätzer
- weitere Modelle
- Vergleich PWM vs. WAM
- Anwendungen
- Zusammenfassung
- Referenzen

# Outline

- 1 Vorbereitungen
- 2 Gliederung
- 3 Inhalte**
- 4 Formulierung
- 5 Standing on the shoulders of giants

# Inhalte

- Abschnitte auf Absatzebene mit Inhalten füllen
- Stichpunkte
- logische Abfolge innerhalb der Abschnitte
- Konsistenz der Inhalte

Wenn man genau weiß, was man aussagen möchte

- schreibt man schneller
- schreibt man exakter
- fallen Formulierungen leichter
- schreibt man mit mehr Begeisterung

# Outline

- 1 Vorbereitungen
- 2 Gliederung
- 3 Inhalte
- 4 Formulierung**
- 5 Standing on the shoulders of giants

# Formulierung

- Formulierungen möglichst exakt, Schwammigkeit vermeiden
- in Texten auch längere Sätze
- Nebensätze sind erlaubt
- auf logischen Aufbau achten und sprachlich deutlich machen
- in deutschen Texten häufig passiv, unpersönlich
- in englischen Texten meist aktiv, auch: unpersönlich mit “one”

# Outline

- 1 Vorbereitungen
- 2 Gliederung
- 3 Inhalte
- 4 Formulierung
- 5 Standing on the shoulders of giants

## Referenzen sind wichtig ...

... um deutlich zu machen,

- was man selbst erarbeitet hat und was nicht
- woher die Inhalte kommen
- evtl. wie die Entwicklung der Erkenntnisse war

... um LeserInnen zu ermöglichen,

- in Originaltexten weiterzulesen
- sich ein "Netzwerk" von Referenzen aufzubauen
- Euren Text als Startpunkt für weitere Recherchen zu nutzen

**So gut wie nie:** Zitate in "" einfach übernehmen

# Wie und wo zitieren?

- Dabei geht es um Inhalte, nicht um Formulierungen (!)
- In Literaturarbeit mindestens nach jedem größeren Absatz
- Alternativ, wenn wenige Quellen: zu Beginn des Abschnittes
- Definitionen, Sätze, . . . , wenn sie nicht von Euch stammen
- Abbildungen, wenn sie nicht von Euch stammen
- Auch sonst alles, wenn es nicht von Euch stammt

# Plain

Wie [3] und [1] zeigen funktioniert die Parametrisierung von [2] so nicht.

## References

- [1] A. Feelders and J. Ivanovs. Discriminative scoring of bayesian network classifiers: a comparative study. In *Proceedings of the third European workshop on probabilistic graphical models*, pages 75–82, 2006.
- [2] Russ Greiner, Xiaoyuan Su, Bin Shen, and Wei Zhou. Structural extension to logistic regression: Discriminative parameter learning of belief net classifiers. *Machine Learning Journal*, 59(3):297–322, 2005.
- [3] Hannes Wettig, Peter Grünwald, Teemu Roos, Petri Myllymäki, and Henry Tirri. On supervised learning of bayesian network parameters. Technical Report HIIT Technical Report 2002-1, Helsinki Institute for Information Technology HIIT, 2002.

# Apalike

Wie [Wettig et al., 2002] und [Feelders and Ivanovs, 2006] zeigen funktioniert die Parametrisierung von [Greiner et al., 2005] so nicht.

## References

- [Feelders and Ivanovs, 2006] Feelders, A. and Ivanovs, J. (2006). Discriminative scoring of bayesian network classifiers: a comparative study. In *Proceedings of the third European workshop on probabilistic graphical models*, pages 75–82.
- [Greiner et al., 2005] Greiner, R., Su, X., Shen, B., and Zhou, W. (2005). Structural extension to logistic regression: Discriminative parameter learning of belief net classifiers. *Machine Learning Journal*, 59(3):297–322.
- [Wettig et al., 2002] Wettig, H., Grünwald, P., Roos, T., Myllymäki, P., and Tirri, H. (2002). On supervised learning of bayesian network parameters. Technical Report HIIT Technical Report 2002-1, Helsinki Institute for Information Technology HIIT.

Wie Wettig et al. [2002] und Feelders and Ivanovs [2006] zeigen funktioniert die Parametrisierung von Greiner et al. [2005] so nicht.

## References

- A. Feelders and J. Ivanovs. Discriminative scoring of bayesian network classifiers: a comparative study. In *Proceedings of the third European workshop on probabilistic graphical models*, pages 75–82, 2006.
- Russ Greiner, Xiaoyuan Su, Bin Shen, and Wei Zhou. Structural extension to logistic regression: Discriminative parameter learning of belief net classifiers. *Machine Learning Journal*, 59(3):297–322, 2005.
- Hannes Wettig, Peter Grünwald, Teemu Roos, Petri Myllymäki, and Henry Tirri. On supervised learning of bayesian network parameters. Technical Report HIIT Technical Report 2002-1, Helsinki Institute for Information Technology HIIT, 2002. URL [citeseer.ist.psu.edu/article/wettig02supervised.html](http://citeseer.ist.psu.edu/article/wettig02supervised.html).

Wie [WGR<sup>+</sup>02] und [FI06] zeigen funktioniert die Parametrisierung von [GSSZ05] so nicht.

## References

- [FI06] A. Feelders und J. Ivanovs. Discriminative Scoring of Bayesian Network Classifiers: a Comparative Study. In *Proceedings of the third European workshop on probabilistic graphical models*, Seiten 75–82, 2006.
- [GSSZ05] Russ Greiner, Xiaoyuan Su, Bin Shen und Wei Zhou. Structural Extension to Logistic Regression: Discriminative Parameter Learning of Belief Net Classifiers. *Machine Learning Journal*, 59(3):297–322, 2005.
- [WGR<sup>+</sup>02] Hannes Wettig, Peter Grünwald, Teemu Roos, Petri Myllymäki und Henry Tirri. On Supervised Learning of Bayesian Network Parameters. Bericht HIIT Technical Report 2002-1, Helsinki Institute for Information Technology HIIT, 2002.