

Filterblasen verstehen

Pöhner, N.
Julius-Maximilians-Universität Würzburg

DOI: 10.18420/ibis-01-01-06

Zusammenfassung

„Vorschläge für dich“ ist ein beliebtes Feature des sozialen Netzwerkes Instagram. Dieses Feature schlägt Usern ähnlichen Inhalte vor, die gut zu ihren bisherigen Interessen passen. Dadurch entstehen sogenannte „Filterblasen“, in denen User personalisierte Feeds erhalten. Soziale Netzwerke nutzen dieses und andere Features, um die User noch länger auf ihrer Plattform zu halten.

Aber wie entstehen die personalisierten Vorschläge? Sie sind das Ergebnis von Data-Mining in sozialen Netzwerken. Dieser Erfahrungsbericht erklärt, wie Data-Mining in sozialen Netzwerken funktioniert, welche Vor- und Nachteile Filterblasen haben und wie diese Inhalte Schülerinnen und Schülern einfach und anschaulich im Informatikunterricht der Sekundarstufe 1 vermittelt werden können.

Einleitung

Die letzte Überarbeitung des Lehrplans für Informatik am bayerischen Gymnasium hat weitreichende inhaltliche Änderungen mit sich gebracht. So fanden insbesondere aktuelle Themen wie Künstliche Intelligenz Einzug in den Lehrplan der Oberstufe (ISB Bayern, 2022). Erste Anknüpfungspunkte an das Thema finden sich allerdings bereits in jüngeren Jahrgangsstufen. So wurde in der 9. Jahrgangsstufe der Lernbereich Datenbanksysteme, der bisher relationale Datenbanksysteme sowie Datenschutz beinhaltete, um das Thema Data-Mining erweitert (ISB Bayern, 2019).

Ein beliebter Ansatz zur Behandlung des Lernbereichs Datenbanksysteme ist die Verwendung von [InstaHub](#), einer Nachbildung des echten sozialen Netzwerkes Instagram zu Lernzwecken. In InstaHub arbeiten Schülerinnen und Schüler als Administratoren des sozialen Netzwerkes. Sie können neue User anlegen, bearbeiten und löschen, per SQL-Datenbankabfragen Informationen über ihre User abfragen sowie personalisierte Werbung schalten. Bisher zeigte sich der Einsatz von InstaHub und der Kontext von sozialen Netzwerken als sehr motivierend und sollte daher auch für das neue Thema Data-Mining als Kontext in meinem Unterricht verwendet werden. Ein Plan für eine Unterrichtseinheit mit vier Doppelstunden zum Thema „Datenschutz und Data-Mining in sozialen Netzwerken“ findet sich im *Extra* am Ende des Artikels.

Soziale Netzwerke wie Instagram versuchen die Inhalte der Feeds ihrer User so auszuwählen, dass sie ihnen gefallen und ihren Interessen entsprechen. Die User liken und teilen diese so häufiger. Die Feeds werden personalisiert und die sozialen Netzwerke erhoffen sich, die User länger auf Ihrer Plattform zu halten. Diese perfektionierten Feeds führen zu sogenannten „Filterblasen“. Auch wenn Filterblasen Usern die Orientierung in der Fülle an Inhalten in sozialen Netzwerken erleichtert, haben diese nicht nur Vorteile. Die User sehen nur Inhalte Gleichgesinnter und müssen sich kaum mit anderen, gegensätzlichen Meinungen auseinandersetzen. Sich eine objektive Meinung zu einem Thema zu bilden, wird so nahezu unmöglich. In politischen Diskussionen kann das zum Beispiel zu einer einseitigen Betrachtung des Sachverhalts führen und befeuert ggf. Desinformation und die Verbreitung von Fake-News. (Bundeszentrale für politische Bildung, 2020, Spiegel Ed, 2023)

Data-Mining „unplugged“

Als Einstieg in das Thema Data-Mining in sozialen Netzwerken bieten sich zwei Aktivitäten ohne den Einsatz eines Computers an. Die erste Aktivität ist eine Aufgabe aus dem Informatikbiber (Bundesweite Informatikwettbewerbe, 2019). Die Schülerinnen und Schüler erlernen bzw. wiederholen mit dieser Aktivität die Datenstruktur des Graphen. Ein Graph besteht aus Knoten und Kanten, die Knoten miteinander verbinden. Im Kontext sozialer Netzwerke repräsentieren

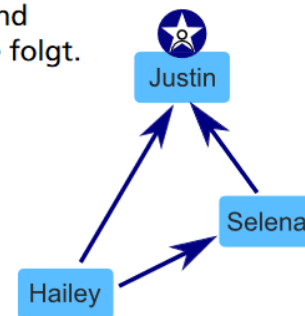
Im sozialen Netzwerk „TeeniGram“ können Mitglieder anderen Mitgliedern „folgen“. In TeeniGram gibt es außerdem Mitglieder-Gruppen. In einer Gruppe nennt man ein Mitglied „Superstar“, wenn

- jedes andere Mitglied der Gruppe ihm folgt und
- es selbst keinem anderen Mitglied der Gruppe folgt.

Ein Beispiel: Eine TeeniGram-Gruppe hat drei Mitglieder, nämlich Hailey, Selena und Justin.

- Hailey folgt Justin und Selena.
- Selena folgt Justin.
- Justin folgt keinem Mitglied der Gruppe.

Justin ist ein Superstar in dieser Gruppe.



Eine andere TeeniGram-Gruppe hat diese Mitglieder: Alan, Don, Frances, Grace und Robin.

- Alan folgt Don und Grace.
- Don folgt Grace und Robin.
- Frances folgt Alan, Grace und Robin.
- Robin folgt Alan und Grace.

Gibt es einen Superstar in dieser Gruppe?

- | | |
|---|---|
| A) Ja, Alan ist ein Superstar in dieser Gruppe. | B) Ja, Frances und Robin sind Superstars in dieser Gruppe. |
| C) Ja, Grace ist ein Superstar in dieser Gruppe. | D) Nein, es gibt keinen Superstar in dieser Gruppe. |

Abbildung 1: Aufgabe "Superstar" aus dem Informatikbiber 2019
(Bildquelle: Bundesweite Informatikwettbewerbe, 2019)

Knoten die Mitglieder eines sozialen Netzwerks und die Kanten geben an, ob ein Mitglied einem anderen „folgt“ bzw. die beiden Mitglieder befreundet sind.

In der zweiten Aktivität lernen die Schülerinnen und Schüler, wie die „Stärke“ einer Freundschaft in sozialen Netzwerken repräsentiert werden kann. Der Jaccard-Index dient dabei als Beispielmaß. Im mathematischen Sinne vergleicht dieser zwei Mengen, zum Beispiel die Interessen zweier Personen, und prüft diese auf Überschneidungen. Je höher der Prozentsatz, den der Jaccard-Index berechnet, desto mehr gemeinsame Interessen haben die beiden Personen. Im Graphen wird dieser Prozentsatz dann als Kantengewicht verwendet.

Zum Beispiel haben die Freunde A und B folgende Interessen:

- Freund A: {Fußball, Gitarre, Computer, Science-Fiction Romane, Brettspiele}
- Freund B: {Klavier, Fußball, Brettspiele, Computer, Hunde}

Der Jaccard-Index berechnet für die Freunde A und B ~42 Prozent gemeinsame Interessen:

$$\frac{| \{ \text{Fußball, Computer, Brettspiele} \} \cap | \{ \text{Fußball, Gitarre, Computer, Science-Fiction Romane, Brettspiele, Klavier, Hunde} \} |}{| \{ \text{Fußball, Gitarre, Computer, Science-Fiction Romane, Brettspiele, Klavier, Hunde} \} |} = \frac{3}{7} = 0.429$$

Bei der Umsetzung im Unterricht werden den Schülerinnen und Schülern zufällig jeweils fünf Interessen zugeteilt (alternativ können diese sie auch selbst wählen). Anschließend laufen sie im Klassenzimmer umher und vergleichen ihre Interessen mit denen ihrer Mitschülerinnen und Mitschüler und berechnen den Anteil gemeinsamer Interessen (ggf. kann auch ein Freundschaftsgraph inklusive ermittelter Kantengewichte erstellt werden).

Soziale Netzwerke analysieren mit Orange3

Die beiden oben genannten Aktivitäten bieten einen ersten Einstieg in die Funktionsweise von Data-Mining in sozialen Netzwerken mit sehr kleiner Datenbasis. Reale Datensätze aus sozialen Netzwerken sind deutlich umfangreicher und lassen sich kaum manuell analysieren. [Orange3¹](#) ist ein mächtiges Werkzeug zur Datenvisualisierung und -analyse, das aber aufgrund seiner intuitiven Benutzeroberfläche auch Einsteigern erlaubt, einfache Datenanalysen durchzuführen. Für den schulischen Einsatz wurde die Verwendung von Orange3 bereits mehrfach vorgeschlagen (zum Beispiel zur Klassifikation mit Entscheidungsbäumen (Grillenberger, Romeike, 2019, Grillenberger, 2021)). Im Unterschied zu diesen Beispielen wird in diesem Artikel mit der Clusteranalyse ein anderes Verfahren von Data-Mining verwendet. Dieses Verfahren wird eingesetzt, um Gruppen von „ähnlichen“ Objekten zu finden.

Die Datenanalyse in Orange3 funktioniert im Allgemeinen durch das Hinzufügen und die Verknüpfung entsprechender Widgets per Drag-and-Drop in der Benutzeroberfläche (für Fortgeschrittene kann Orange3 auch als Python-Bibliothek genutzt werden). Zur Analyse

¹ <https://orangedatamining.com>



Name des Widgets	Network File	Network Explorer	Network Clustering
Bild	 Network File	 Network Explorer	 Network Clustering
Beschreibung	Liest ein Netzwerk aus einer Datei ein.	Stellt das Netzwerk grafisch dar.	Findet Cluster in einem Netzwerk.

Tabelle 1: Widgets zur Clusteranalyse in Orange3

sozialer Netzwerke bietet Orange3 das Add-on „Network analysis“². Die für uns wichtigen Widgets sind in Tabelle 1 zu sehen.

Die Schülerinnen und Schüler nutzen diese Widgets, um Filterblasen in sozialen Netzwerken zu entdecken. Die Network Files stellen dabei einfache Excel-Dateien dar, die Freundschafterpaare und die Stärke der Freundschaft angeben. So können soziale Netzwerke als gewichteter Graph repräsentiert werden. Im Beispieldatensatz wurde mithilfe der Excel-Funktion ZUFALLSBEREICH(<Untere_Zahl>;<Obere_Zahl>) zufällige Kantengewichte im Graphen generiert. Die User wurden aus InstaHub exportiert. Das Network File kann den Schülerinnen und Schülern je nach Vorwissen entweder vorgegeben oder von ihnen selbst mithilfe entsprechender SQL-Datenbankabfragen in InstaHub und der genannten Excel-Funktion erstellt werden (falls vorher ein eigener Freundschaftsgraph in der oben genannten Aktivität erstellt wurde, kann auch dieser als Startpunkt genutzt werden).

Abbildungen 2 und 3 zeigen die Datenanalyse mit Orange3 sowie das Ergebnis der Clusteranalyse. Ein „Cluster“ beschreibt dabei eine Gruppe von ähnlichen Objekten. Diese Objekte sind meist durch mehrdimensionale Attribute gekennzeichnet, auf Basis derer sie gruppiert werden können. Die Ähnlichkeit wird dabei mit einem bestimmten Ähnlichkeitsmaß (zum Beispiel dem Euklidischen Abstand) ermittelt. In einem sozialen Netzwerk können neben gleichen Interessen auch weitere Merkmale wie gemeinsame Freunde, gelikte Inhalte, etc. zu hoher Ähnlichkeit beitragen, auf Basis derer sie bei einer Clusteranalyse zusammengefasst werden.

Im [Onlinematerial](#)³ sind neben dem Beispieldatensatz auch das zugehörige Orange3-Projekt zu finden.

² <https://orangedatamining.com/widget-catalog/networks/networkfile/>

³ https://github.com/nicolaipoehner/orange3_socialnetworkanalysis

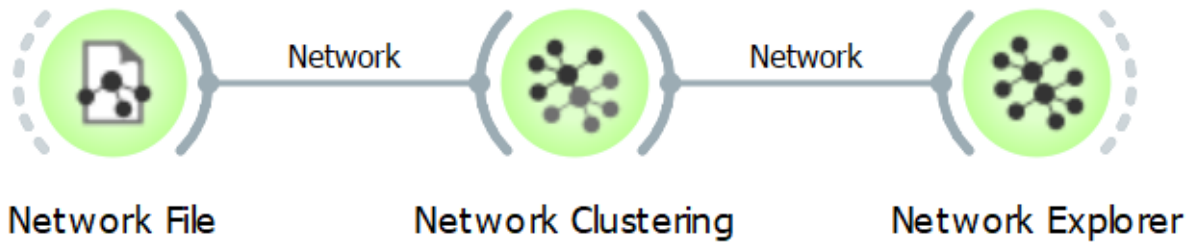


Abbildung 2: Clusteranalyse mit Orange3

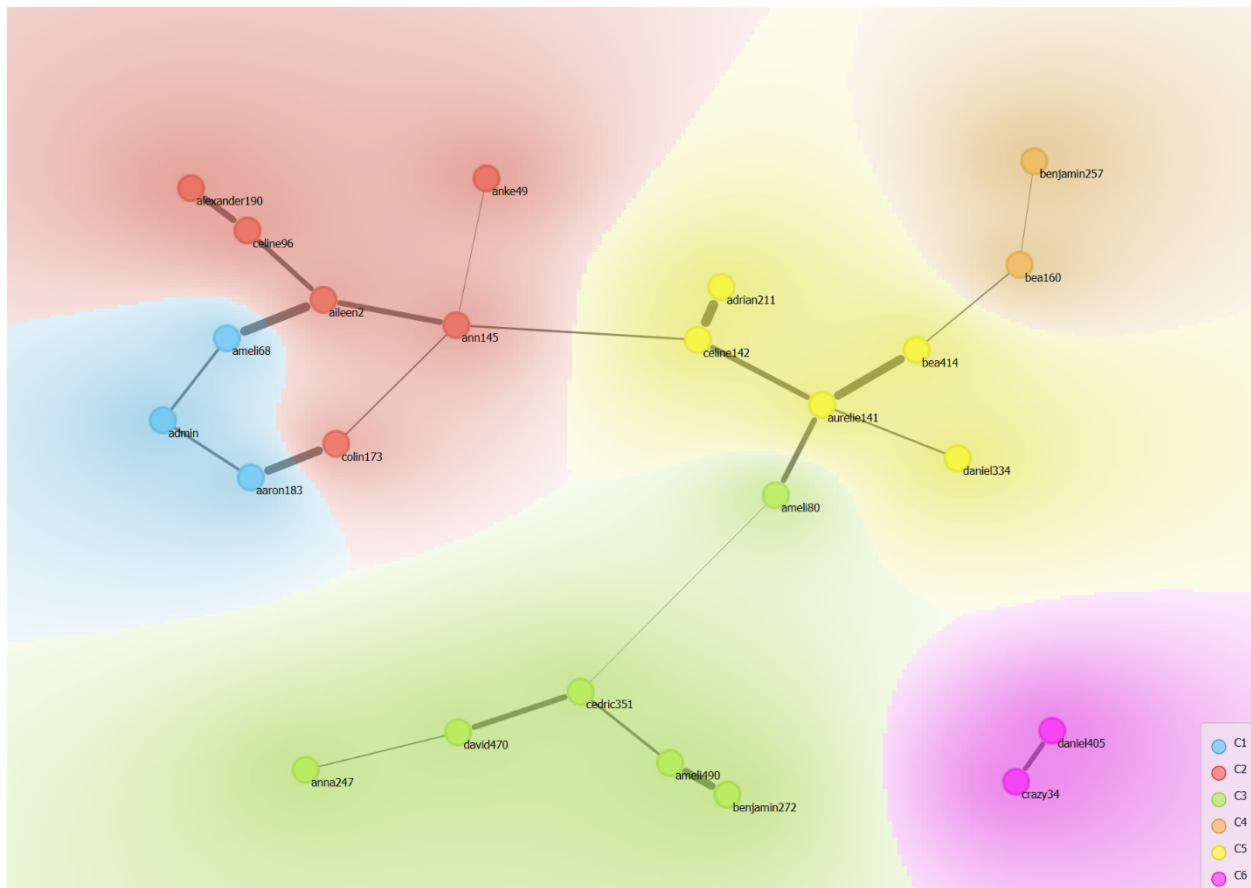


Abbildung 3: Ergebnis der Clusteranalyse mit Orange3

Was können wir gegen Filterblasen tun?

Abbildung 3 zeigt das Ergebnis der Clusteranalyse. Aufgrund ihrer Ähnlichkeit wurden die User also in mehrere Cluster zusammengefasst. Diese entsprechen ihren Filterblasen.

Die Behandlung von Data-Mining im Kontext sozialer Netzwerke erlaubt neben einer technologischen (*Wie entstehen Filterblasen?*) und gesellschaftlich-kulturellen Perspektive

(Welche positiven und negativen Effekte haben Filterblasen?) ebenfalls die Betrachtung der anwendungsbezogenen Perspektive im Sinne des Dagstuhl-Dreiecks (Gesellschaft für Informatik, 2016). Die anwendungsbezogene Perspektive beantwortet dabei die Frage „Was kann ich (als User) tun, um die Effekte von Filterblasen zu minimieren?“ In Instagram können User zum Beispiel die vorgeschlagenen Inhalte deaktivieren⁴, um selbst über die Inhalte ihres Feeds zu bestimmen. Den Schülerinnen und Schülern können also konkrete Handlungsempfehlungen an die Hand gegeben werden.

Extra: Plan für eine Unterrichtseinheit mit vier Doppelstunden zum Thema „Datenschutz und Data-Mining in sozialen Netzwerken“

1. Einführung in das Thema Datenschutz

Zum Einstieg in das Thema sehen die Schülerinnen und Schüler das Video „Warum weiß Facebook, dass ich schwanger bin“⁵. Sie diskutieren das Video und berichten von persönlichen Erfahrungen bzgl. Datenschutz in sozialen Netzwerken. Anschließend erarbeiten sie sich eine Definition der Begriffe „personenbezogene Daten“ und „Datenschutz“. Sie lernen außerdem den rechtlichen Rahmen von Datenschutz in Deutschland kennen.

2. Datenschutz in practice: Views und Benutzerrechte

Nach der Einführung der theoretischen Grundlagen erarbeiten die Schülerinnen und Schüler, wie Datenschutz in sozialen Netzwerken konkret umgesetzt werden kann. Mithilfe von Benutzerrechten und Views erstellen sie verschiedene Ansichten für Profile in sozialen Netzwerken, zum Beispiel für öffentliche und private Profile in InstaHub.

3. Data-Mining in sozialen Netzwerken, Filterblasen

Die Schülerinnen und Schüler betrachten das Feature „Vorschläge für dich“ in Instagram näher und diskutieren dessen Funktionsweise. Sie erarbeiten sich den Begriff „Filterblase“ und erklären, wie das Feature „Vorschläge für dich“ zu Filterblasen führen kann. Dazu nutzen sie die beiden genannten „unplugged“ Aktivitäten. Sie informieren sich anschließend anhand des Videos „Filterbubble – Das Leben in der Informationsblase“⁶ über Filterblasen sowie deren Vor- und Nachteile.

⁴ <https://about.instagram.com/de-de/blog/tips-and-tricks/control-your-instagram-feed>

⁵ <https://www1.wdr.de/mediathek/av/video-warum-facebook-weiss-dass-du-schwanger-bist--102.html>

⁶ <https://www.bpb.de/lernen/bewegt-bild-und-politische-bildung/webvideo/fakefilter/315166/folge-08-bubble-bubble-filterbubble/>

4. Data-Mining in practice: Soziale Netzwerke analysieren mit Orange3

Zum Einstieg sehen die Schülerinnen und Schüler den satirischen Beitrag über Filterblasen in sozialen Netzwerken „[Filterblase Facebook](#)“⁷. Danach analysieren sie mithilfe von Orange3 Beispieldatensätze von sozialen Netzwerken.

Fazit

Der Kontext von Filterblasen in sozialen Netzwerken erwies sich als sehr motivierend zur Behandlung des Thema Data-Mining für die Schülerinnen und Schüler in meinem Unterricht, da sie alle schon eigene Erfahrungen mit Filterblasen gesammelt hatten. Die persönlichen Erfahrungen der Schülerinnen und Schüler zeigen aber (leider) auch die große Relevanz des Themas, da viele von ihnen sich oft der negativen Konsequenzen ihrer schlecht konfigurierten Einstellung zur Privatsphäre nicht bewusst waren oder sie sogar (zugunsten bequemer Features des sozialen Netzwerks wie dem automatischen Anmelden) bewusst ignorierten.

Quellen

Alle Webseiten/Links wurden zuletzt geprüft am 01.05.2023.

Bundesweite Informatikwettbewerbe (BWIInf) (2019). Informatikbiber: Aufgaben 2019. URL: <https://bwinf.de/fileadmin/biber/2019/Biberheft2019-mit-Loesungen.pdf>

Bundeszentrale für politische Bildung (bpb) (2020). Fakefilter: Folge 08: Bubble Bubble Filterbubble: Die Filterblase – alles nur im Kopf?. URL: <https://www.bpb.de/lernen/bewegt-bild-und-politische-bildung/webvideo/fakefilter/315166/folge-08-bubble-bubble-filterbubble/>.

Gesellschaft für Informatik (GI) (2016). Dagstuhl-Erklärung: Bildung in der digitalen Welt. URL: <https://dagstuhl.gi.de/dagstuhl-erklaerung>.

Grillenberger, A., Romeike, R. (2019). About Classes and Trees: Introducing Secondary School Students to Aspects of Data Mining. In: Pozdniakov S., Dagienė V. (eds.) „Informatics in Schools. New Ideas in School Informatics. ISSEP 2019.“ Cham: Springer. URL: https://link.springer.com/chapter/10.1007/978-3-030-33759-9_12.

Grillenberger, A. (2021). Datenflussorientierte Big-Data-Analyse (mit Orange3). In: LOG IN. 195/196. S. 61-66. URL: <https://dl.gi.de/handle/20.500.12116/36195>.

ISB Bayern (2019). LehrplanPLUS – Gymnasium – Informatik 9. URL: <https://www.lehrplanplus.bayern.de/fachlehrplan/gymnasium/9/informatik>

ISB Bayern (2022). LehrplanPLUS – Gymnasium – Informatik 11 (NTG). URL: <https://www.lehrplanplus.bayern.de/fachlehrplan/gymnasium/11/informatik/ntg>

⁷ <https://www.ardmediathek.de/video/extra-3/filterblase-facebook/das-erste/Y3JpZDovL25kci5kZS9hZWY4ZGI0OS1hMDMyLTQ2NzQtYTlhOS1jMDI1MDI1NWQ4ZTE>

Spiegel Ed (2023). Unterrichtsmaterial zu Fake News und Filterblase. URL:
<https://ed.spiegel.de/unterrichtsmaterial/fake-news-und-filterblase>

Lizenz



Dieser Artikel steht unter der Lizenz CC BY-NC-SA 4.0 zur Verfügung.

Kontakt

Dr. Nicolai Pöhner

Julius-Maximilians-Universität Würzburg

Didaktik der Informatik

nicolai.poehner@uni-wuerzburg.de