

# Selektives Lernen für Empfehlungsmaschinen<sup>1</sup>

Pawel Matuszyk<sup>2</sup>

## Abstract:

Das Problem der Informationsüberladung ist im digitalen Zeitalter besonders herausfordernd. Um Nutzer beim Finden relevanter Artikel zu unterstützen, wurden Empfehlungsalgorithmen entwickelt. Diese Algorithmen lernen Nutzerpräferenzen und sagen voraus, was Nutzer in der Zukunft relevant finden werden. Empfehlungsmaschinen lernen aus historischen Daten zum Nutzerfeedback. Diese Daten sind jedoch von fehlenden Werten dominiert, da jeder Nutzer Feedback zu nur wenigen Produkten bereitstellen kann. Das hat zur Folge, dass konventionelle Ansätze zum Lernen von Nutzerpräferenzen *alle verfügbaren Daten* verwenden. In dieser Arbeit schlagen wir *selektives Lernen* vor. In diesem *alternativen Ansatz* werden *Daten, Nutzer, oder Aspekte der Modelle nur selektiv zum Trainieren der Präferenzmodelle* verwendet. Wir zeigen, dass dieser Ansatz zu einer *signifikanten Verbesserung der Empfehlungsqualität* führt.

## 1 Einführung

Empfehlungsmaschinen (recommender systems) mildern das Problem der Informationsüberladung, das immer dann auftritt, wenn Menschen aus einer Vielzahl von Alternativen wählen müssen, von denen nur wenige relevant sind. Heutzutage sind Menschen mit diesem Problem häufiger denn je konfrontiert (z.B. bei einer Vielzahl von Filmen, Büchern, Musikstücken oder Medikamenten). Die Anwendungen der Empfehlungsmaschinen umfassen jedoch nicht nur E-Commerce [Li14, SKR99, PNH15], sondern auch Medizin [WP14, Ch12], Lehrmaterialien [Ma11, WLZ15], u.v.a.m.

Empfehlungsmaschinen lernen Nutzerpräferenzen aus historischem Feedback und sagen voraus, welche Artikel für einen gegebenen Nutzer in der Zukunft relevant sein werden. Die relevanten Artikel werden dann individuell einem gegebenen Empfänger empfohlen. Das historische Feedback von Nutzern ist allerdings oft sehr spärlich, da jeder Nutzer nur relativ wenige Artikel wahrnehmen kann. Dies führt zu einer häufigen Annahme in der Forschung zu Empfehlungsmaschinen, dass alle verfügbaren Daten zum Lernen der Präferenzmodelle genutzt werden sollen. In diesem Beitrag und in der zugrundeliegenden Dissertation schlagen wir ein neues Paradigma vor, laut dem die Trainingsdaten vorsichtig selektiert werden sollten. Wir bezeichnen das neue Paradigma als *selektives Lernen*. Wir entwickeln drei Typen von Ansätzen zum selektiven Lernen, sowohl für strombasierte, als auch für batch-basierte Algorithmen.

Die strom-basierten Algorithmen, die im Fokus dieser Arbeit liegen, haben gegenüber den batch-basierten Algorithmen einen entscheidenden Vorteil. Sie sind adaptiv und können

---

<sup>1</sup> Englischer Titel der Dissertation: "Selective Learning for Recommender Systems" [Ma17a]

<sup>2</sup> Fakultät für Informatik, Otto-von-Guericke-Universität Magdeburg, Deutschland, pawel.matuszyk@ovgu.de

neue Information sofort in ihre Präferenzmodelle integrieren. Somit können sie, idealerweise in Echtzeit, auf Änderungen reagieren.

Unser erster Ansatz zum selektiven Lernen sind Vergessensmethoden für strombasierte Empfehlungsmaschinen. Sie selektieren, welche Daten vergessen werden sollten. Somit entscheiden sie auch, welche Daten zum Lernen der Modelle verwendet werden. Die Vergessensmethoden stellen eine weitere Möglichkeit zur Adaptation der Modelle an Änderungen der Nutzerpräferenzen dar. Wir betonen, dass nicht nur veraltete Daten vergessen werden können. Wir schlagen elf Vergessensstrategien vor, die die zu vergessende Informationen selektieren, und drei alternative Algorithmen, die das Vergessen umsetzen.

Unser zweiter selektiver Ansatz eignet sich für Selektion der Nachbarn in Collaborative-Filtering-Algorithmen (CF). Die CF-Algorithmen arbeiten ähnlich wie k-Nearest-Neighbours im maschinellen Lernen. Sie suchen nach ähnlichen Nutzern und empfehlen die von ihnen als relevant empfundenen Produkte. Unser selektives Kriterium erlaubt die Selektion der Nachbarn nicht nur aufgrund der Ähnlichkeit, sondern auch aufgrund der Zuverlässigkeit der Information.

Unser letzter selektiver Ansatz basiert auf teil-überwachtem Lernen. Wir haben das erste teil-überwachte Framework für strom-basierte Empfehlungsmaschinen entwickelt. Dieses Framework erlaubt es den Empfehlungsmaschinen, von den umfangreichen fehlenden Daten zu lernen. Das ist z.B. im Co-Training-Ansatz möglich, wo mehrere Algorithmen parallel arbeiten und sich gegenseitig trainieren, indem sie eigene Vorhersagen anderen Algorithmen als Trainingsbeispiele zur Verfügung stellen. Nicht alle Vorhersagen allerdings vertrauenswürdig. Das Lernen aus falschen Vorhersagen könnte eine Verschlechterung der Empfehlungen zur Folge haben. Um das zu verhindern, schlagen wir weitere Selektionsmechanismen vor, die es erlauben, nur aus zuverlässigen Vorhersagen zu lernen.

Unsere Evaluierung auf realen Daten zeigt, dass selektives Lernen eine wesentliche Verbesserung der Qualität der Empfehlungen im Vergleich zu Systemen ohne selektives Lernen mit sich bringt.

## 2 Selektives Lernen für Empfehlungsmaschinen

In diesem Kapitel definieren wir das selektive Lernen für Empfehlungsmaschinen und schlagen drei Ansätze vor, die diese Definition auf unterschiedliche Arten implementieren.

**Definition:** *Selektives Lernen* für Empfehlungsmaschinen umfasst Methoden zum Lernen und Vorhersagen von Nutzerpräferenzen nicht aufgrund aller verfügbaren Daten, sondern aufgrund einer Selektion von Daten und Aspekten der Modelle, die die Qualität der Empfehlungen maximieren.

Eine formale Definition befindet sich in der Dissertation, die die Grundlage für diesen Beitrag ist [Ma17a]. Sie kann hier aus Platzgründen nicht ausführlich beschrieben werden.

## 2.1 Selektives Vergessen

Selektives Vergessen erfordert zwei grundlegende Neuerungen bei der strom-basierten Verarbeitung von Ratings (eine Form von Nutzerfeedback). Diese Neuerungen spiegeln sich in den Komponenten einer Empfehlungsmaschine wieder. Die Abbildung 1 bietet eine Übersicht über die Komponenten und die Architektur einer solchen Empfehlungsmaschine.

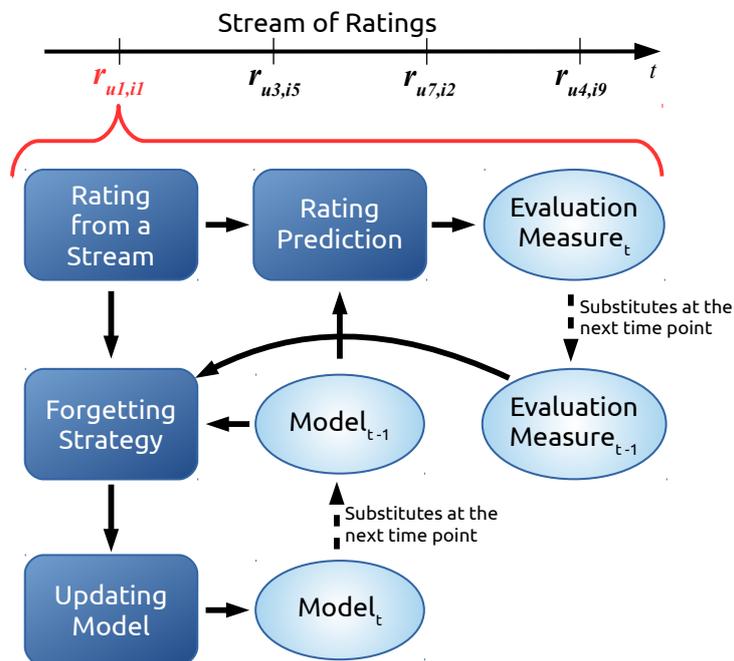


Abb. 1: Übersicht über die Komponenten einer strom-basierten Empfehlungsmaschine mit selektivem Vergessen [Ma17a]

Die erste der neuartigen Komponenten ist eine Vergessensstrategie. Diese Komponente entscheidet, welche Informationen vergessen werden sollen. Dabei muss beachtet werden, dass nicht nur alte Informationen vergessen werden können. Im Kapitel 2.1.1 geben wir ein Beispiel für eine solche Strategie.

Die zweite neuartige Komponente setzt das Vergessen der Information um, die durch eine Vergessensstrategie selektiert wurde. Ein einfaches Löschen der Daten aus einer Datenbank ist nicht ausreichend, da ein Modell bereits von den zu löschenden Daten gelernt hat. Diese Komponente entfernt den Einfluss der zu löschenden Daten aus dem Modell. Sie ist in dem Update-Mechanismus der strom-basierten Algorithmen angesiedelt (s. Abbildung 1). Wir haben drei alternative Algorithmen zur Umsetzung des Vergessens vorgeschlagen und ausführlich evaluiert. Dabei haben wir einen Repräsentanten der Matrix-Faktorisierung-Algorithmen, den BRISMF Algorithmus [Ta09], mit unseren Vergessensmethoden erweitert.

### 2.1.1 Vergessensstrategien

Im Rahmen der Dissertation wurden elf Vergessensstrategien vorgeschlagen, die einer der zwei Kategorien zugeordnet werden können:

1. Rating-basierte Vergessensstrategien
2. Vergessensstrategien im latenten Raum

Aus Platzgründen geben wir hier jedoch nur ein Beispiel für eine solche Vergessensstrategie.

**Sensitivitätsbasiertes Vergessen** Diese Strategie ist ein Beispiel für eine rating-basierte Vergessensstrategie. Sie arbeitet direkt mit Ratings (Nutzerfeedback) und entscheidet für jedes Rating, ob es vergessen oder behalten werden sollte.

Diese Vergessensstrategie basiert auf der Idee der lokalen Sensitivität der Präferenzmodelle. Sobald ein neues Rating beobachtet wurde, verwenden strom-basierte Algorithmen dieses Rating zur Aktualisierung des zugehörigen Nutzerprofils. Infolge dieser Aktualisierung wird das Profil angepasst. Der Grad der Veränderung des Profils kann gemessen werden.

Wenn das gegebene Rating konsistent mit dem bisherigen Nutzerprofil war, dann soll es nur eine geringfügige Anpassung des Modells hervorrufen. Wenn eine große Änderung des Profils beobachtet wurde, dann kann angenommen werden, dass das Rating nicht in die bisherigen Präferenzen des Nutzers passt. Das betroffene Rating könnte ein Ausreißer sein (z.B. ein Geschenk für eine andere Person).

Um solche Ausreißer zu erkennen, kann der Grad der lokalen Veränderung des Nutzermodells gemessen werden. In Matrix-Faktorisierung-Algorithmen, die die State-of-the-art in Empfehlungsmaschinen sind, wird ein Nutzermodell in Form eines Vektors mit latenten Faktoren gespeichert (s. [Ta09] für die Erklärung der Grundlagen zum latenten Raum). Sei  $p_u^t$  ein Vektor des Nutzers  $u$  zum Zeitpunkt  $t$ . Mit der folgenden Variable  $\Delta_{p_u}$  kann der Unterschied im Modell eines Nutzers zwischen den Zeitpunkten  $t$  und  $t + 1$  im  $k$ -dimensionalen latenten Raum erfasst werden.

$$\Delta_{p_u} = \sum_{i=0}^k (p_{u,i}^{t+1} - p_{u,i}^t)^2 \quad (1)$$

$\Delta_{p_u}$  kann über die Zeit, im Verlauf eines Datenstroms beobachtet werden. Somit können der Mittelwert  $\overline{\Delta_{p_u}}$  und Standardabweichung  $SD(\Delta_{p_u})$  ermittelt werden. Eine Veränderung des Nutzermodells gilt dann als abnormal hoch (z.B. bei einem Ausreißer), wenn die folgende Ungleichung gilt:

$$\Delta_{p_u} > \overline{\Delta_{p_u}} + \alpha \cdot SD(\Delta_{p_u}) \quad (2)$$

$\alpha$  ist ein Hyperparameter, der die Sensitivität der Vergessenstrategie kontrolliert. Er wird experimentell so bestimmt, dass die Empfehlungsqualität maximal ist.

## 2.1.2 Ergebnisse zum Selektiven Vergessen

Um den Einfluss von Vergessensmethoden auf die Empfehlungsqualität zu untersuchen, haben wir mehr als 1040 Experimente durchgeführt. Dabei haben wir eine State-of-the-art-Methode mit selektivem Lernen erweitert und haben dann die Erweiterung mit der Standard-Variante verglichen. Die Experimente umfassten eine Offline-Evaluierung mit Testdaten, Hyperparameteroptimierung und statistischen Tests mit dem Friedman-Test und Wilcoxon-Rangsummentest. Die Ergebnisse wurden gegen multiples Testen mit Hommel's Methode korrigiert, um die Alphafehler-Kumulierung zu vermeiden. Als Evaluierungsmaß haben wir das inkrementelle Recall von Cremonesi et al. verwendet [CKT10].

Die Abbildung 2 zeigt eine Auswahl der Ergebnisse unserer Evaluierung (für alle Ergebnisse s. [Ma17a]). Links in den Plots ist die Vergleichsmethode (Matrixfaktorisierung ohne selektives Lernen). Die sonstigen Boxplots zeigen die Ergebnisse der Matrixfaktorisierung mit unterschiedlichen Vergessensstrategien. In unseren Experimenten haben wir eine *signifikante Verbesserung* der Empfehlungsqualität auf sieben von acht Datensätzen beobachtet.

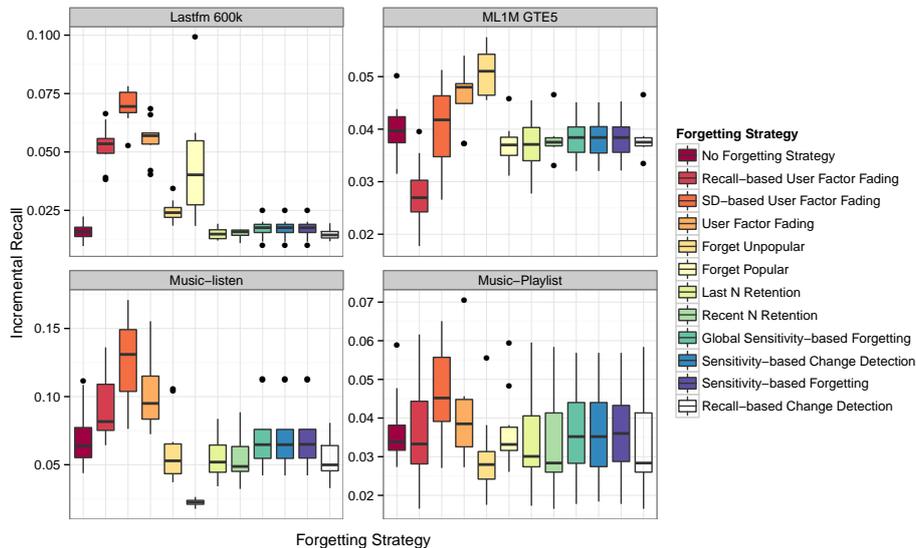


Abb. 2: Inkrementelles Recall einer Empfehlungsmaschine mit und ohne Vergessensmethoden auf vier Datensätzen mit implizitem Nutzerfeedback. Auf jedem Datensatz gibt es Vergessensstrategien, die die Vergleichsbaseline (No Forgetting) übertreffen (höhere Werte sind besser) [Ma17b].

## 2.2 Hoeffding-CF

Unser zweite Ansatz zum selektiven Lernen wurde für nachbarschaftsbasierte Collaborative Filtering Methoden (CF) entwickelt. Die CF-Methoden funktionieren auf eine ähnliche Art und Weise wie die konventionellen kNN-Methoden (k Nearest Neighbours) im maschinellen Lernen. Vereinfacht ausgedrückt, suchen sie nach ähnlichen Nutzern und empfehlen, was die ähnlichen Nutzer relevant fanden. Diese Methoden arbeiten oft mit einem Schwellenwert für die Ähnlichkeit zwischen potenziellen Nachbarn und dem aktiven Nutzer.

Wir schlagen ein weiteres Kriterium für eine mehr restriktive Selektion der Nachbarn vor. Dieses Kriterium basiert auf Hoeffding-Bound, einer theoretischen Schranke für eine Abweichung einer Beobachtung vom Mittelwert. Unser Kriterium berücksichtigt nicht nur die absolute Ähnlichkeit zwischen zwei Nutzern, sondern auch die Anzahl der Beobachtungen, die zur Berechnung dieser Ähnlichkeit verwendet wurden. Somit, berücksichtigt unser Kriterium die Zuverlässigkeit von Ähnlichkeiten.

Laut dem von uns vorgeschlagenen Kriterium ist der aktive Nutzer  $u_a$  nur dann zu einem weiteren Nutzer  $u_x$  zuverlässig ähnlich, wenn die folgende Ungleichung gilt:

$$\widehat{sim}(u_a, u_x) - \widehat{sim}(u_a, u_B) > 2\varepsilon \quad (3)$$

$u_B$  ist ein sogenannter Baseline-Nutzer. Für Baseline-Nutzer haben wir mehrere Implementierungsmöglichkeiten vorgeschlagen (s. [Ma17a]). Intuitiv kann man sie als Durchschnittsnutzer oder zufällige Nutzer verstehen. Das bedeutet, dass die Ähnlichkeit zwischen  $u_a$  und  $u_x$  nur dann zuverlässig ist, wenn sie signifikant höher ist als die Ähnlichkeit von  $u_a$  zum Baseline-Nutzer. Der Begriff der Signifikanz und der Wert von  $\varepsilon$  wurden von der Hoeffding-Bound hergeleitet. Demnach gilt [Ho63]:

$$\varepsilon = \sqrt{\frac{R^2 \cdot \ln(1/\delta)}{2n}} \quad (4)$$

wo  $n$  die Anzahl der Beobachtungen ist, die zur Berechnung der Ähnlichkeit verwendet wurden.  $R$  steht für den Wertebereich der Ratings und  $\delta$  bestimmt das Signifikanzniveau.

### 2.2.1 Ergebnisse zu Hoeffding-CF

Unsere Experimente auf vier reellen Datensätzen zeigen, dass unser selektives Kriterium verwandte Methoden (wie Shrinkage [BKV07] oder Significance Weighting [He99]) übertrifft. Eine grafische Darstellung der Ergebnisse ist in der Abbildung 3 zu sehen.

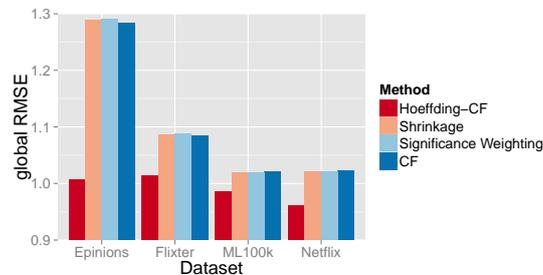


Abb. 3: Ergebnisse der CF-Methoden mit unterschiedlichen Methoden der Nachbar-Selektion [MS14]

### 2.3 Teil-überwachtes Lernen

Unser letzte Ansatz zum selektiven Lernen ist teil-überwachtes Lernen für strom-basierte Empfehlungsmaschinen. In diesem Ansatz lernen Empfehlungsmaschinen von den zahlreichen fehlenden Werten in der Rating-Matrix (nicht bewertete Artikel). Da die Anzahl der fehlenden Werte in der Matrix extrem hoch ist (typischerweise 99%), müssen die Methoden, die diese Information nutzen, selektiv vorgehen. Wir schlagen das erste Framework für teil-überwachtes Lernen für strom-basierte Empfehlungsmaschinen vor.

In unserem Framework vergleichen wir zwei Ansätze zum teil-überwachten Lernen: Co-Training und Self-Learning. Im Co-Training werden mehrere Modelle parallel trainiert. Die Vorhersage eines Modells kann dann als Trainingsbeispiel durch die restlichen Modelle verwendet werden. Auf diese Weise trainieren sich die Modelle gegenseitig, indem sie Vorhersagen füreinander als Trainingsbeispiele bereitstellen. Im Self-Learning ist das Prinzip ähnlich. Anstatt jedoch aus Vorhersagen anderer Modelle zu lernen, lernt das Modell in diesem Fall aus eigenen Vorhersagen.

Das Lernen aus Vorhersagen bringt allerdings Risiken mit sich. Wenn die Vorhersagen inkorrekt sind, dann lernen die Modelle aus falschen Daten. Um dem Problem vorzubeugen, dürfen die Modelle nicht aus allen Vorhersagen lernen, sondern sie handeln *selektiv*. Wir schlagen zahlreiche Methoden zur Auswahl der zuverlässigen Vorhersagen vor, aus denen die Modelle anschließend lernen.

Um das selektive teil-überwachte Lernen auf Datenströmen umzusetzen, haben wir auch weitere neuartigen Komponenten in unserem Framework eingeführt. Die Abbildung 4 stellt eine Übersicht dieser Komponenten dar. Eine der wichtigsten Komponenten ist das Zuverlässigkeitsmaß. Dieses Maß wird genutzt, um die Zuverlässigkeit einer Vorhersage zu schätzen. Diese ist von essenzieller Bedeutung bei der Entscheidung, ob ein Modell aus einer bestimmten Vorhersage lernen soll, oder nicht.

Eine weitere wichtige Komponente ist das Training-Set-Splitter. Diese Komponente sorgt dafür, dass die Modelle, die parallel trainiert werden, unterschiedliche Daten bekommen. Wäre das nicht der Fall, dann wären die Modelle identisch und sie könnten sich gegenseitig

nichts beibringen. Diese und weitere neuartige Komponenten des Frameworks werden im Details in der Dissertationsschrift beschrieben [Ma17a].

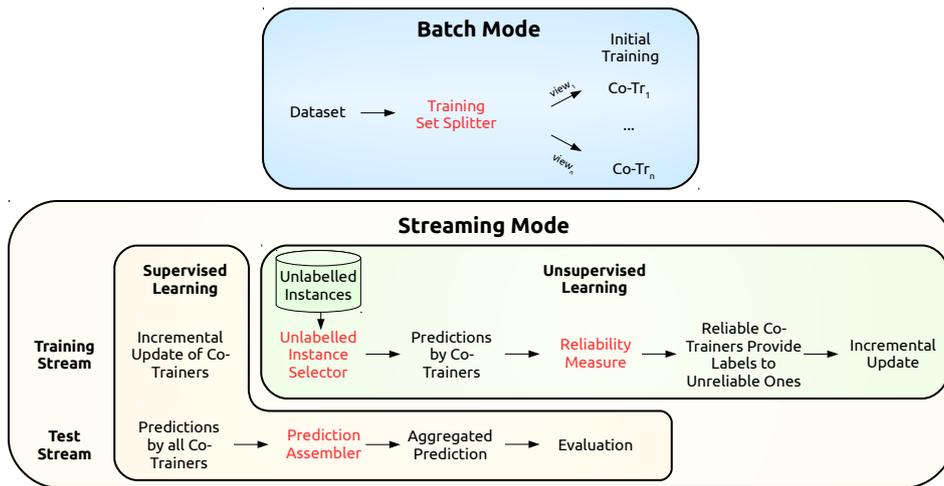


Abb. 4: Übersicht der Komponenten des Frameworks für teil-überwachtes Lernen für strom-basierte Empfehlungsmaschinen [MS17]

### 2.3.1 Ergebnisse zum Teil-überwachten Lernen

Um das selektive Lernen aus Vorhersagen inklusive der zahlreichen Komponenten des neuartigen Frameworks zu evaluieren, haben wir mehr als 700 Experimente durchgeführt. In diesen Experimenten haben wir Co-Training mit zwei (SSL2) und mit drei (SSL3) parallelen Modellen mit Self-Learning (SL) und mit einer Baseline ohne teil-überwachtes Lernen (NoSSL) verglichen.

Die Abbildung 5 stellt die Ergebnisse auf einem der getesteten Datensätze dar. Das Evaluierungsmaß ist das früher erwähnte inkrementelle Recall (höhere Werte sind besser). Aus der Abbildung ist ersichtlich, dass Co-Training mit drei Modellen (SSL3) alle anderen Methoden dominiert. Experimente auf anderen Datensätzen haben ähnliche Ergebnisse geliefert. Unsere statistische Analyse hat ergeben, dass SSL3 auf allen getesteten Datensätzen die Empfehlungsqualität gegenüber NoSSL signifikant verbessert hat. Dies geht jedoch mit einem erhöhten Rechenaufwand einher.

## 3 Zusammenfassung

In dieser Arbeit haben wir eine neues Lern-Paradigma für Empfehlungsmaschinen, *selektives Lernen*, vorgeschlagen. Entgegen der geltenden Annahme, dass alle verfügbaren Daten zum Lernen genutzt werden sollen, haben wir gezeigt, dass eine intelligente Auswahl der Lerndaten zu signifikant besseren Empfehlungen führt.

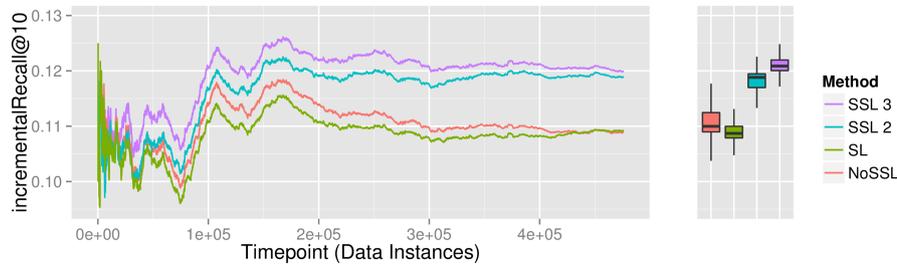


Abb. 5: Inkrementelles Recall der teil-überwachten Methoden (SSL und SL) und der Basline ohne teil-überwachtes Lernen (NoSSL) auf dem Movielens 1M Datensatz (höhere Werte sind besser) [MS17].

Wie haben drei Arten des selektiven Lernens entwickelt: selektives Vergessen von Rating oder Modellteilen, Selektion der Nachbarn im nachbarschaftsbasierten Collaborative Filtering, und Selektion der Vorhersagen für teil-überwachtes Lernen von Nutzerpräferenzen. Bei allen diesen Methoden des selektiven Lernens hat unsere ausführliche Evaluierung mit reellen Daten gezeigt, dass das neue Paradigma die Qualität der Empfehlungen *signifikant verbessert*.

In der Zukunft kann des Weiteren erforscht werden, wie sich die Kombination von mehreren selektiven Methoden in einem System auf die Performance der Empfehlungsmaschinen auswirkt. Da vor Allem das teil-überwachte Lernen zu längeren Rechenzeiten führt, kann in der Zukunft das Potenzial der verteilten Berechnung einzelner Modelle auf unterschiedlichen Rechnern untersucht werden.

## Literaturverzeichnis

- [BKV07] Bell, Robert; Koren, Yehuda; Volinsky, Chris: Modeling relationships at multiple scales to improve accuracy of large recommender systems. In: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 2007.
- [Ch12] Chen, Rung-Ching; Huang, Yun-Hou; Bau, Cho-Tsan; Chen, Shyi-Ming: A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection. Expert Systems with Applications, 39(4):3995–4006, 2012.
- [CKT10] Cremonesi, Paolo; Koren, Yehuda; Turrin, Roberto: Performance of Recommender Algorithms on Top-n Recommendation Tasks. In: Proceedings of ACM RecSys. RecSys '10. ACM, S. 39–46, 2010.
- [He99] Herlocker, Jonathan L.; Konstan, Joseph A.; Borchers, Al; Riedl, John: An algorithmic framework for performing collaborative filtering. In: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, New York, NY, USA, S. 230–237, 1999.
- [Ho63] Hoeffding, Wassily: Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc., 58:13–30, 1963.

- [Li14] Lin, Zhijie: An empirical investigation of user and system recommendations in e-commerce. *Decision Support Systems*, 68:111–124, 2014.
- [Ma11] Manouselis, Nikos; Drachler, Hendrik; Vuorikari, Riina; Hummel, Hans; Koper, Rob: Recommender Systems in Technology Enhanced Learning. In (Ricci, Francesco; Rokach, Lior; Shapira, Bracha; Kantor, Paul B., Hrsg.): *Recommender Systems Handbook*. Springer US, Boston, MA, S. 387–415, 2011.
- [Ma17a] Matuszyk, Pawel: *Selective Learning for Recommender Systems*. Dissertation, 2017.
- [Ma17b] Matuszyk, Pawel; Vinagre, João; Spiliopoulou, Myra; Jorge, Alípio Mário; Gama, João: Forgetting techniques for stream-based matrix factorization in recommender systems. *Knowledge and Information Systems*, Aug 2017.
- [MS14] Matuszyk, Pawel; Spiliopoulou, Myra: Hoeffding-CF: Neighbourhood-Based Recommendations on Reliably Similar Users. In (Dimitrova, Vania; Kuflik, Tsvi; Chin, David; Ricci, Francesco; Dolog, Peter; Houben, Geert-Jan, Hrsg.): *User Modeling, Adaptation, and Personalization*, Jgg. 8538 in *Lecture Notes in Computer Science*, S. 146–157. Springer International Publishing, 2014.
- [MS17] Matuszyk, Pawel; Spiliopoulou, Myra: Stream-based semi-supervised learning for recommender systems. *Machine Learning*, S. 1–28, 2017.
- [PNH15] Paraschakis, D.; Nilsson, B. J.; Holländer, J.: Comparative Evaluation of Top-N Recommenders in e-Commerce: An Industrial Perspective. In: *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*. S. 1024–1031, Dec 2015.
- [SKR99] Schafer, J.; Konstan, J.; Reidl, J.: *Recommender Systems in E-Commerce*. In: *Proceedings of ACM Conference on Electronic Commerce*. Denver, Colorado, USA, November 1999.
- [Ta09] Takács, Gábor; Pilászy, István; Németh, Bottyán; Tikk, Domonkos: Scalable Collaborative Filtering Approaches for Large Recommender Systems. *J. Mach. Learn. Res.*, 10, 2009.
- [WLZ15] Wu, Dianshuang; Lu, Jie; Zhang, Guangquan: A fuzzy tree matching-based personalized e-learning recommender system. *IEEE Transactions on Fuzzy Systems*, 23(6):2412–2426, 2015.
- [WP14] Wiesner Martin; Pfeifer Daniel: *Health Recommender Systems: Concepts, Requirements, Technical Basics and Challenges*. *International Journal of Environmental Research and Public Health*, 11(3):2580–2607, feb 2014.



**Pawel Matuszyk** ist ein wissenschaftlicher Mitarbeiter am Lehrstuhl für Knowledge Management and Discovery an der Otto-von-Guericke-Universität in Magdeburg. Er promovierte in 2017 auf dem Gebiet von Machine Learning und Data Science. Zu seinen Forschungsschwerpunkten gehören adaptive Empfehlungsmaschinen, Stream-Mining, teilüberwachtes Lernen und inkrementelle Matrix-Faktorisierung. Seine Master- und Bachelor-Abschlüsse erhielt er von der Otto-von-Guericke-Universität in Wirtschaftsinformatik.