DEEPLINK TECHNISCHER BERICHT · NOVEMBER 2026

Share of Voice und Zuspruch in der Xiaohongshu-Suche: Messmethode

Measuring Share of Voice and Group Affinity in Xiaohongshu Search

DeepLink International, Data Science
深通（杭州）国际教育科技有限公司 · Hangzhou, China

Zusammenfassung Chinesische Kursteilnehmer informieren sich bei der Wahl einer deutschen Sprachschule vor allem über die Suchergebnisse auf Xiaohongshu und die Kommentare darunter. Die Suchergebnisse unterscheiden sich je nach Zeitpunkt und Nutzer, eine einzelne Beobachtung ist nicht reproduzierbar; zudem ist die Wirkung einzelner Beiträge und Kommentare sehr ungleich verteilt. Dieser Bericht beschreibt ein wiederholbares Messverfahren: Die Stabilität der Stichprobe wird über die Split-Half-Konsistenz mehrerer Suchdurchläufe geprüft, das Gewicht eines Beitrags aus Rangdiskontierung und Interaktion bestimmt. Darauf bauen drei Kennzahlen auf: Informationsanteil, Share of Voice und Zuspruch. Alle drei sind beschränkte, dimensionslose Anteile, unempfindlich gegenüber Mehrfachkommentaren und einzelnen viralen Beiträgen und über Suchbegriffe und Zeitpunkte hinweg vergleichbar.

Schlagwörter Share of Voice; Zuspruch; Split-Half-Konsistenz; Rangdiskontierung; Suche in sozialen Medien

---

1 Fragestellung und Notation

Ein Suchbegriff wird in K Durchläufen abgefragt; je Durchlauf werden die ersten 60 Ergebnisse erfasst. Der Rang des Beitrags n im Durchlauf r sei $p_{n,r}$; erscheint der Beitrag nicht, ist seine Rangdiskontierung 0. Die stabil erscheinenden, thematisch relevanten Beiträge bilden die Stichprobe N; die Teilmenge der von eigenen Accounts veröffentlichten Beiträge sei N°. Die Kommentare unter Beitrag n bilden die Menge C_n. Ein besprochenes Objekt (eine Schule) wird mit b bezeichnet. Jede Nennung einer Schule trägt eine Polarität $\sigma \in [-1, 1]$: −1 steht für klare Ablehnung, 1 für klare Zustimmung.

2 Stabilität der Stichprobe

Die Reihenfolge der Suchergebnisse enthält einen Zufallsanteil. Zuerst ist deshalb zu prüfen, ob die beobachtete Rangfolge stabil ist. Wir übernehmen den Gedanken der Split-Half-Reliabilität aus der Testtheorie [2, 3]: Die k Durchläufe werden zufällig in zwei gleich große Hälften A und B geteilt. Für jede Hälfte werden nach der kumulierten Rangdiskontierung die 20 obersten Beiträge $T_{20}(A)$ und $T_{20}(B)$ bestimmt. Ihr Überlappungsanteil ist die Konsistenz einer Teilung:

$$O(A,B) = \frac{|T_{20}(A) \cap T_{20}(B)|}{\min(\,|\,T_{20}(A)\,|\,,\,|\,T_{20}(B)\,|\,)} \tag{1}$$

Gemittelt wird über $S = 60$ unabhängige Zufallsteilungen (fester Startwert, daher reproduzierbar):

$$C_k = \frac{1}{S}\sum_{j=1}^{S} O(A_j, B_j) \tag{2}$$

Ab dem 6. Durchlauf wird C_k alle 2 Durchläufe berechnet. Liegt die Spannweite der letzten 4 Werte nicht über der Toleranz $\varepsilon = 0{,}04$, gilt die Stichprobe als konvergiert und die Erhebung endet; mindestens sind also 12 Durchläufe nötig:

$$\max_{i \in I_k} C_i - \min_{i \in I_k} C_i \le \varepsilon, \quad I_k = \{k-6, k-4, k-2, k\} \tag{3}$$

Suchbegriffe mit einem Konvergenzwert unter 0,60 werden als „streuend“ gekennzeichnet; ihre Kennzahlen dienen nur der Orientierung. Für nicht konvergierte Erhebungen werden keine Kennzahlen berechnet.

3 Gewicht eines Beitrags

Je weiter hinten ein Beitrag steht, desto seltener wird er gesehen. Wir verwenden die in der Evaluation von Suchsystemen übliche logarithmische Rangdiskontierung [1]:

$$g(p) = \frac{1}{\log_2(p+1)} \tag{4}$$

Die Exposition eines Beitrags ist der Mittelwert seiner Rangdiskontierung über alle Durchläufe:

$$E_n = \frac{1}{K}\sum_{r=1}^{K} g(p_{n,r}) \tag{5}$$

Der Interaktionswert gewichtet Likes L_n, Speicherungen F_n und Kommentare M_n; maßgeblich ist der jeweils letzte Messwert der Erhebung:

$$H_n = 10L_n + 15F_n + 20M_n \tag{6}$$

Das Gewicht eines Beitrags ist das Produkt aus Exposition und der Wurzel des Interaktionswerts. Durch die Wurzel wächst das Gewicht unterproportional zur Interaktion, sodass einzelne virale Beiträge das Ergebnis nicht dominieren:

$$w_n = E_n\sqrt{H_n} \tag{7}$$

4 Sichtbarkeit eines Kommentars

Die Wirkung eines Kommentars hängt davon ab, wie oft er gesehen wird. Kommentar c habe ℓ_c Likes und sei die j_c-te Antwort in seinem Thread (für Kommentare erster Ebene gilt $j_c = 0$). Die ersten beiden Antworten eines Threads werden nicht abgewertet, weitere mit ihrer Position zunehmend:

$$v_c = \frac{1 + \ell_c}{1 + 0{,}5 \cdot \max(0, j_c - 1)}, \quad s_c = \sqrt{v_c} \tag{8}$$

Alle weiteren Berechnungen verwenden s_c, aus demselben Grund wie beim Gewicht eines Beitrags.

5 Definition der Kennzahlen

5.1 Informationsanteil

Der Informationsanteil misst, welcher Teil der Inhalte in den Suchergebnissen von eigenen Accounts stammt:

$$R = \frac{\sum_{n \in N^\circ} w_n}{\sum_{n \in N} w_n} \tag{9}$$

5.2 Share of Voice

Share of Voice erfasst nur, wie oft eine Schule genannt wird, unabhängig von der Wertung. Nennt der Text eines Beitrags die Schule b, erhält b das volle Gewicht des Beitrags. Nennt ein Kommentar b, erhält b den Anteil, der der Sichtbarkeit dieses Kommentars an allen Kommentaren des Beitrags entspricht:

$$W_b = \sum_{n \in N} w_n \left(\mathbb{1}_{b \in n} + \frac{1}{D_n} \sum_{c \in C_n(b)} s_c \right) \tag{10}$$

Dabei ist $C_n(b)$ die Menge der Kommentare unter Beitrag n, die b nennen, und D_n die normierende Summe der Kommentarsichtbarkeit des Beitrags. Um Mehrfachkommentare zu dämpfen, gilt: Hat ein Nutzer unter einem Beitrag mindestens 3 Kommentare geschrieben und stammen mindestens 20 % der Kommentare von ihm, werden seine Nennungen derselben Schule zu einer zusammengefasst; es zählt das größte s_c, und auch in D_n geht nur dieser Höchstwert ein. Share of Voice ist der normierte Anteil:

$$V_b = \frac{W_b}{\sum_{b'} W_{b'}} \tag{11}$$

Offensichtlich gilt $V_b \in [0,1]$ und $\sum_b V_b = 1$.

5.3 Zuspruch

Der Zuspruch stützt sich nur auf Kommentare und misst, wem die Kommentierenden insgesamt zustimmen. Zunächst wird je Beitrag gerechnet. Das Stimmgewicht einer kommentierenden Person u ist ihr sichtbarster Kommentar unter dem Beitrag, normiert über alle Kommentierenden:

$$\pi_{u,n} = \frac{\max_{c \in C_n(u)} s_c}{\sum_{u'} \max_{c \in C_n(u')} s_c} \tag{12}$$

Mehrere Äußerungen derselben Person zur selben Schule werden gemittelt zu $\bar{\sigma}_{u,b,n}$. Der Nettozuspruch der Schule b in diesem Beitrag ist:

$$\nu_{b,n} = \sum_u \pi_{u,n} \bar{\sigma}_{u,b,n} \tag{13}$$

Der Eindruck, den ein Beitrag hinterlässt, wird in der Regel von der Mehrheitsmeinung in den Kommentaren bestimmt. Der Nettowert wird deshalb mit dem Exponenten $\gamma = 3$ geschärft, sodass die führende Schule einen größeren Anteil erhält; das Vorzeichen bleibt erhalten:

$$a_{b,n} = \operatorname{sgn}(\nu_{b,n}) \frac{|\nu_{b,n}|^\gamma}{\sum_{b'} |\nu_{b',n}|^\gamma} \tag{14}$$

Zwei Korrekturen kommen hinzu. Die Äußerungsdichte δ_n gibt an, wie eindeutig die Kommentare unter dem Beitrag Stellung beziehen. Der Personenfaktor $\varphi_{b,n}$ wertet Schulen ab, die nur von einzelnen Personen genannt werden; $U_{b,n}$ ist die Zahl der verschiedenen Personen, die b unter dem Beitrag nennen, der Schwellenwert ist $U_0 = 3$:

$$\delta_n = \min\left(1, \sum_b |\, \nu_{b,n} \,|\right), \quad \varphi_{b,n} = \min\left(1, \frac{U_{b,n}}{U_0}\right) \tag{15}$$

Über die Beiträge wird mit dem Beitragsgewicht summiert und mit der Summe der Beträge aller Schulen normiert:

$$A_b = \sum_{n \in N} w_n a_{b,n} \delta_n \varphi_{b,n} \tag{16}$$

$$G_b = \frac{A_b}{\sum_{b'} |A_{b'}|} \tag{17}$$

Damit gilt $G_b \in [-1, 1]$ und $\sum_b |G_b| = 1$. $G_b < 0$ bedeutet, dass ablehnende Äußerungen überwiegen.

6 Eigenschaften

Beschränkt und vergleichbar. Alle drei Kennzahlen sind Anteile und hängen nicht von der absoluten Größe der Stichprobe ab. Sie lassen sich über Suchbegriffe und Erhebungen hinweg vergleichen.

Robust gegenüber Extremwerten. Interaktion und Sichtbarkeit gehen mit der Wurzel ein; gewichtet wird je Person, nicht je Kommentar; wiederholte Nennungen von Vielkommentierenden werden zusammengefasst. Ein einzelner viraler Beitrag oder ein einzelner Nutzer kann das Ergebnis nicht bestimmen.

Reproduzierbar. Die Berechnung der Kennzahlen enthält keinen Zufallsanteil; die Split-Half-Prüfung verwendet einen festen Startwert. Dieselben Daten ergeben stets dasselbe Ergebnis.

Gleicher Maßstab für alle. Für Partnerschulen und alle anderen Schulen gelten dieselben Formeln und Parameter.

7 Parameter und Stichprobe

Tabelle 1 Parameterwerte

Symbol	Bedeutung	Wert
—	erfasste Ergebnisse je Durchlauf	60
T_{20}	Länge der Liste im Split-Half-Vergleich	20
S	Zahl der Zufallsteilungen	60
ε	Konvergenztoleranz (Spannweite der letzten 4 Werte)	0,04
—	Gewichte der Interaktion (Likes / gespeichert / Kommentare)	10 / 15 / 20
—	Abwertung von Antworten im Thread	0,5
γ	Schärfungsexponent des Zuspruchs	3
U_0	Schwellenwert des Personenfaktors	3
—	Kriterium für Vielkommentierende	≥ 3 Kommentare und ≥ 20 %

Tabelle 2 Stichproben der auf dieser Website verwendeten Erhebungen (8. und 9. November 2026)

Suchbegriff	Suchläufe	Konsistenz	Beiträge	Kommentare	Informationsanteil
Sprachkurs Deutschland 德国语言班	31,5	77,4 %	38	664	51,6 %
Sprachschule Deutschland 德国语言学校	38,8	67,0 %	35	533	39,1 %
Sprachkurs Essen 埃森语言班	17,5	79,3 %	26	423	37,6 %
Sprachkurs Deutschland Kosten 德国语言班费用	15,8	74,3 %	38	538	33,2 %

8 Grenzen

Erstens stammen die Daten nur von einer Plattform, Xiaohongshu. Zweitens ist jede Erhebung eine Momentaufnahme und kein langfristiger Durchschnitt. Drittens ist die Stichprobe begrenzt: je Suchbegriff etwa 30 bis 40 Beiträge und einige hundert Kommentare; weiter hinten platzierte Schulen werden selten genannt, ihr Rang ist nicht stabil. Viertens enthält der Share of Voice der Partnerschulen auch Inhalte eigener Accounts; zum Vergleich berechnen wir zusätzlich den Anteil nur aus Inhalten Dritter. Fünftens werden Schulnamen so erfasst, wie sie in den Kommentaren tatsächlich genannt werden; sie können vom offiziellen Namen abweichen.

Literatur

1. Järvelin, K., & Kekäläinen, J. (2002). Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems*, 20(4), 422–446.
2. Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3(3), 271–295.
3. Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3(3), 296–322.

© 2026 深通（杭州）国际教育科技有限公司 www.deeplink-edu.com