

Applying Voting Methods in User Research

Likert versus Condorcet: Kann man Wahlverfahren verwenden um Konsens bei Nutzerpräferenzen zu ermitteln?

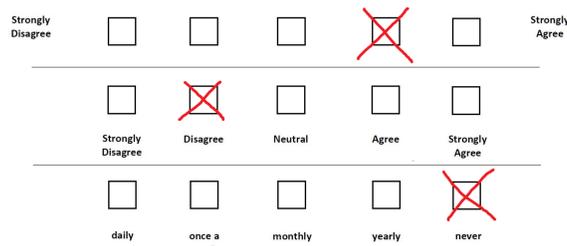
Erheben von Nutzerkonsens? Status Quo: Likert-Skala*

Ziel: Für individuell unterschiedlich bewertete Eigenschaften von Benutzerschnittstellen muss man einen Konsens finden, um die "beste" Lösung für eine Mehrheit von Nutzern zu wählen.

Beispiele: Farbe, Größe, Gewicht eines Mobiltelefons, Dauer einer Animation.

Status Quo: Ausfüllen von *Likert-type Items** zur Bewertung der einzelnen Alternativen. Punktwerte pro Option, Mittelwert bilden.

*Likert-Skalen (intervallskalierte Daten) bestehen aus mehreren Likert-type Items (ordinalskalierte Daten), die kalibriert verrechnet werden.



Likert-type Items: a) mit Extremen, b) mit Labels für jede Antwortoption oder c) mit eindeutig unterschiedlichen Abständen zwischen Antwortoptionen.

Likert-type Items

- Abfragen von Einschätzungen
- weit verbreitet
- eigentlich ordinalskaliert
- absolute Bewertung jeder einzelnen präsentierten Alternative

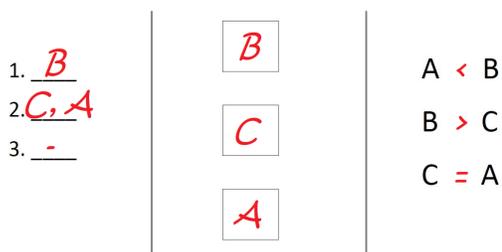
Likert-Probleme: Oft keine geeichten Skalen. Zuordnung Option ↔ Punktzahl willkürlich. Auswertung der Ergebnisse durch deskriptive Statistik (Mittelwert, etc.) wird kritisch gesehen.

Frage: Kann man auch konservativere Wahlverfahren aus der Social-Choice-Theorie nutzen, um einen "besten" Kandidaten ermitteln.

Erhebung von Präferenzen mittels Ranking/Reihung von Alternativen.

Ermitteln konsensfähiger Alternativen mittels einer Condorcet-Methode.

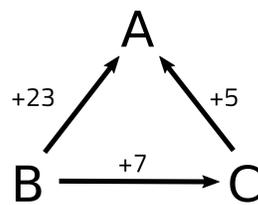
Mögliche Alternativen: Ranking und Wahlverfahren (Condorcet und Borda)



Ranking-Möglichkeiten: eigene Reihung, anordnen von Karten mit Alternativen, paarweise Vergleiche.

Ranking von Alternativen

- ordinalskaliert
- Ausdruck der relativen Präferenz von mehreren präsentierten Alternativen
- Auszuwerten durch Wahlverfahren
- diverse Präsentationsmöglichkeiten, auch paarweise Vergleiche möglich.



Graph der paarweisen Präferenzen basierend auf mehreren Einzel-Rankings. B gewinnt Wahl vor C und A. Condorcet-Methoden lösen auch Patts auf.

Condorcet-Methoden

- Vorzugswahlverfahren; der Kandidat gewinnt, der jedem anderen im Vergleich vorgezogen wurde
- konsens-basiert
- Marquis de Condorcet (1785)

Schulze-Methode (1997):

- weit verbreitete Condorcet-Methode
- Bestimmen eines Gewinners mittels Graph der Präferenzen

Simple Alternative: Borda Count

- Pro Ranking erhält jede Alternative Punkte entsprechend erreichtem Rang
- Alternative mit den meisten Punkten in Summe gewinnt.

Vorteile:

- robuste, gut untersuchte Methoden für ordinale Daten
- Ranking auch paarweise möglich ("dieses ist besser als das vorherige")

Pilot-Studie: bevorzugte Mauszeigergeschwindigkeit

Motivation: Studie, um zu ermitteln wie gut verschiedene Verfahren zum Erheben von Nutzerkonsens auch die objektiv beste Option ermitteln:
- Erhebung: Likert vs. Ranking
- Auswertung: Mittelwert, Median, Schulze, Borda Count.

Teilnehmer: 24 (14 m), Alter 20-59

Ablauf: Klassischer Pointing-Test mit einer Computermaus. Vier den TN unbekannte Zeigergeschwindigkeiten. Mehrfache Wiederholung. Objektive Performance beim Treffen der Ziele wird als Ground Truth gemessen.

Danach: präferierte Rangfolge bilden und jede Geschwindigkeit einzeln auf einer Skala bewerten.

Pointing-Test: Erst den Go-Button anklicken und danach möglichst schnell den blauen Kreis.

Auswertung: objektive Ground Truth und subjektive Bewertungen auf den Skalen wurden mithilfe deskriptiver Statistik, der Schulze- und der Borda-Methode ausgewertet. Rankings wurden nur mit Schulze-Methode und Borda-Count ausgewertet.

Ranking und Wahlverfahren liefern ähnliche Ergebnisse wie Skalenbewertung

	Ground Truth				Likert				Ranking	
	Median	Mean	Schulze	Borda	Median	Mean	Schulze	Borda	Schulze	Borda
1.	B	B	B	B	BC	B	B	B	B	B
2.	A	A	C	C	BC	C	C	C	C	C
3.	C	C	A	A	AD	AD	A	A	A	A
4.	D	D	D	D	AD	AD	D	D	D	D

Gesamtranking aller Nutzer für vier Mauszeigergeschwindigkeiten A(0,5), B(0,875), C(1,5), D(3,0)

Ergebnisse: Gewinner und Verlierer bei allen Varianten identisch zu Ground Truth. Komplette Reihung unabhängig von Erhebungsmethode identisch, wenn Wahlverfahren zur Auswertung verwendet werden.

Diskussion: Studie mit explorativem Charakter (simples Setup, geringe Zahl an präsentierten Alternativen, keine Replikation)

Ranking zur Präferenzhebung und die Wahlverfahren zur Auswertung individueller Rankings können valide und robuste Alternativen zu Likert-Skalen und Mittelwerten sein.

Weitere Informationen unter:

https://hci.ur.de/projects/voting_methods

Gefördert vom Bayerischen Staatsministerium für Wissenschaft und Kunst im Rahmen des