

Accelerating iteration speed: ranker evaluation with debiased interleaving

„Full assortment“ two-sided marketplace

GMV 7 billion EUR

Up to **35** orders per second

12.2 million active customers

18 million products



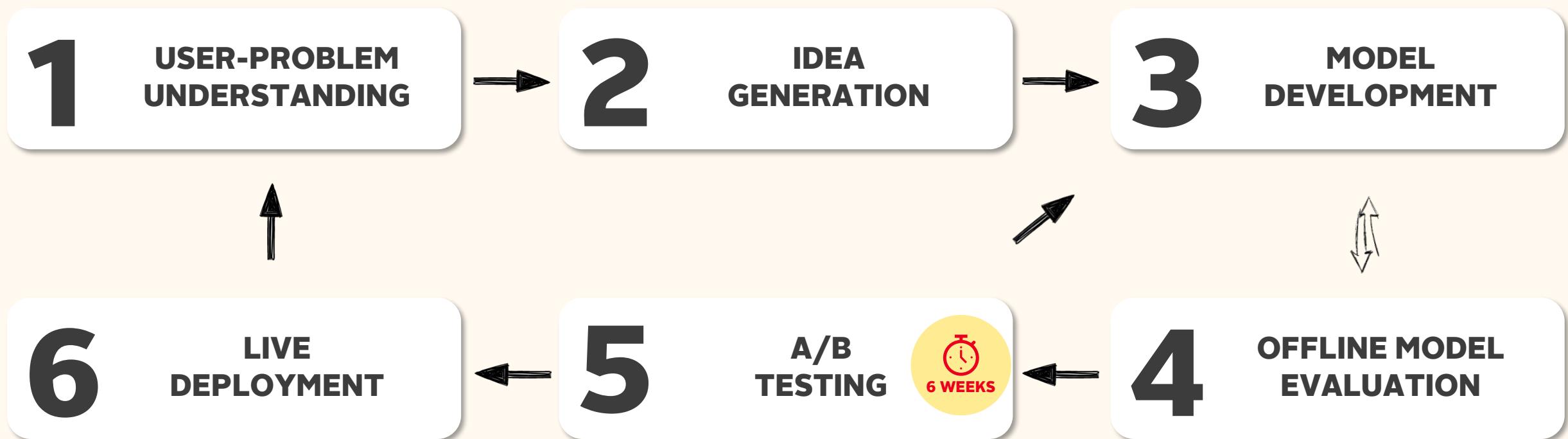
Dr. Leonie Brinkmann
Senior Digital Analyst - Ranking

Mika Wolle
Product Manager - Ranking

“Product development is a risky business. If you don’t evolve your product quickly enough, competitors will outrun you.”¹



A/B test runtime is a problem for our iteration speed



In the literature we found a promising approach

Debiased Balanced Interleaving at Amazon Search

Nan Bi

Amazon, Palo Alto, CA, USA

Pablo Castells

Amazon, Madrid, Spain

Daniel Gilbert

Amazon, Palo Alto, CA, USA

Beyond A/B Test : Speeding up Airbnb Search Ranking Experimentation through Interleaving



Qing Zhang · Follow

Published in The Airbnb Tech Blog · 10 min read · Oct 6, 2022

Innovating Faster on Personalization Algorithms at Netflix Using Interleaving

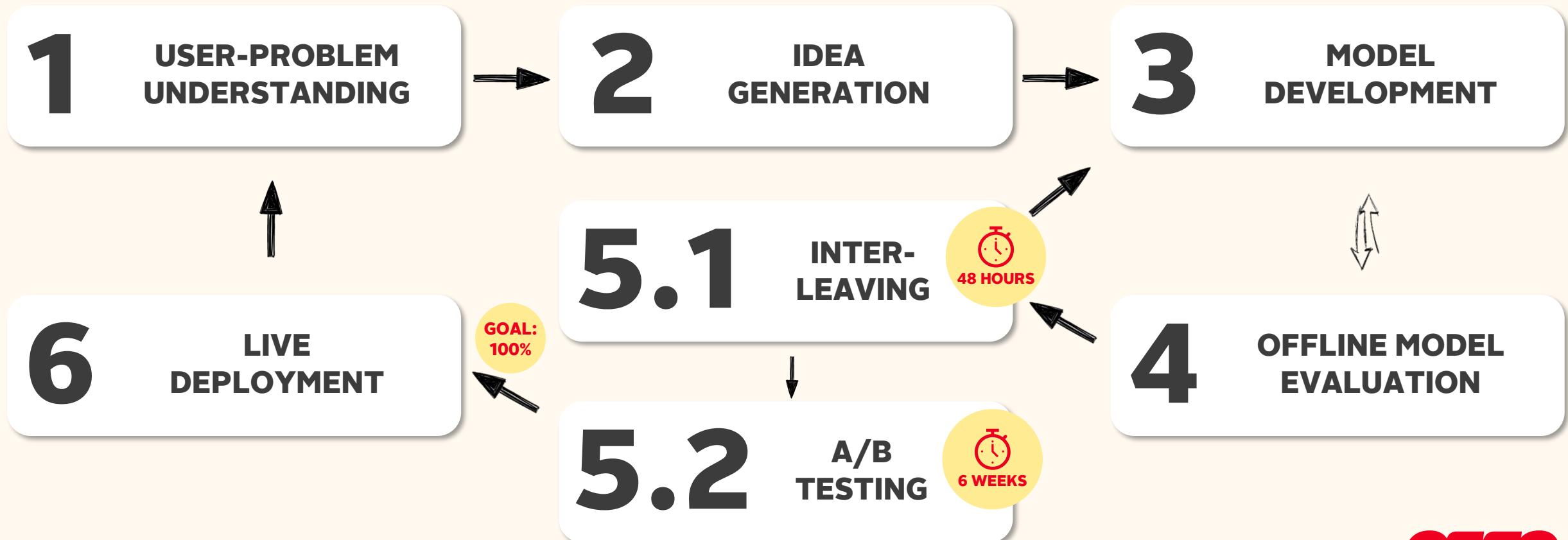


Netflix Technology Blog · Follow

Published in Netflix TechBlog · 8 min read · Nov 29, 2017

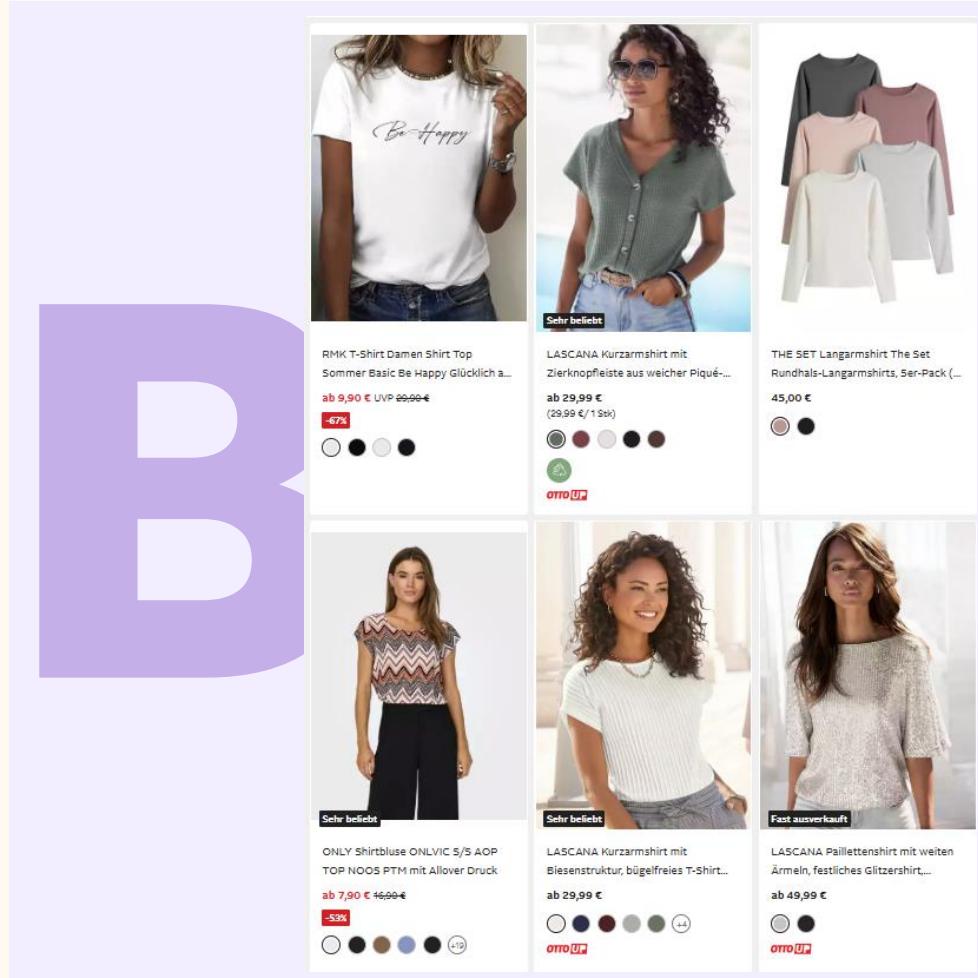
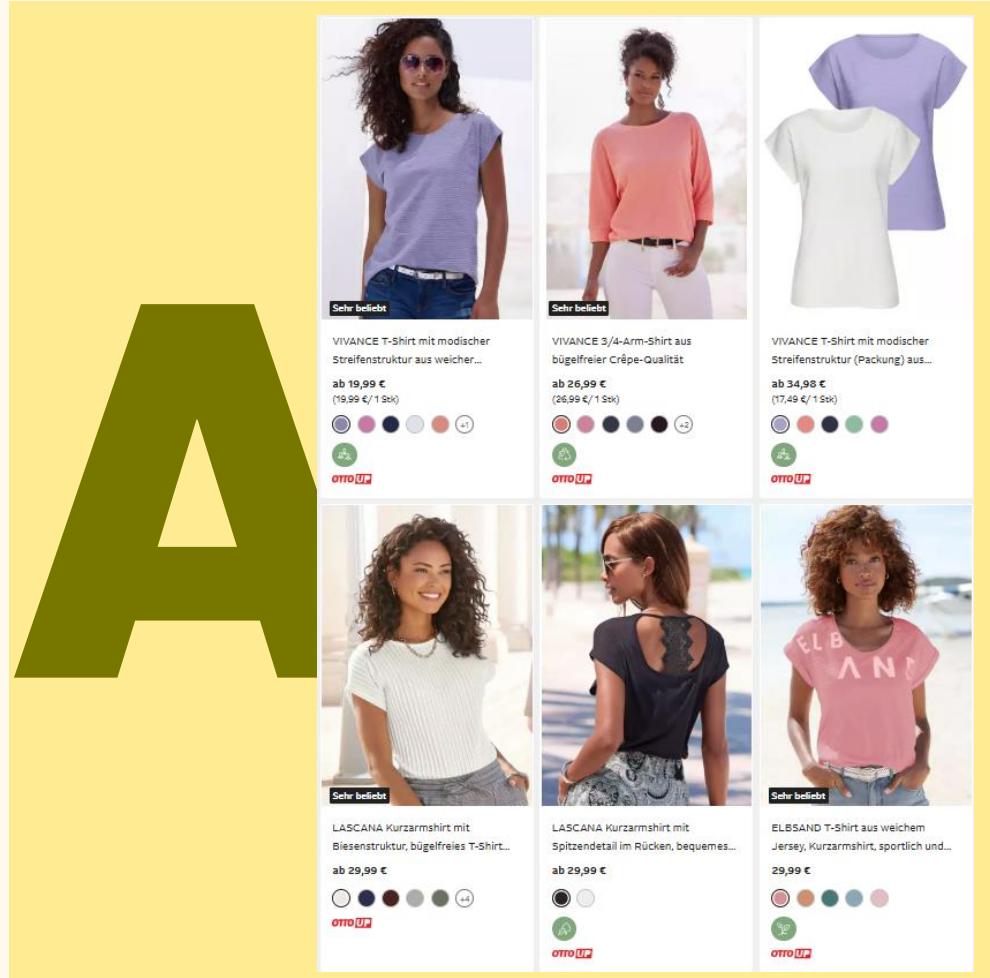
OTTO

Interleaving enables us to collect meaningful feedback within 48 hours



In a classic A/B test users either see ranking A or B

A



OTTO

With interleaving both rankers are presented to the same user simultaneously

A

Sehr beliebt

VIVANCE T-Shirt mit modischer Streifenstruktur aus weicher...
ab 19,99 € (19,99 €/1 Stück)
OTTO

Sehr beliebt

VIVANCE 3/4-Arm-Shirt aus bügelfreier Crêpe-Qualität
ab 26,99 € (26,99 €/1 Stück)
OTTO

Sehr beliebt

VIVANCE T-Shirt mit modischer Streifenstruktur (Packung) aus...
ab 34,98 € (17,49 €/1 Stück)
OTTO

Sehr beliebt

LASCANA Kurzarmshirt mit Biesenstruktur, bügelfreies T-Shirt...
ab 29,99 €
OTTO

Sehr beliebt

LASCANA Kurzarmshirt mit Spitzendetail im Rücken, bequemes...
ab 29,99 €
OTTO

Sehr beliebt

ELBSAND T-Shirt aus weichem Jersey, Kurzarmshirt, sportlich und...
29,99 €
OTTO

B

Sehr beliebt

RMK T-Shirt Damen Shirt Top Sommer Basic Be Happy Glücklich a...
ab 9,90 € UVP 29,99 €
-67%
OTTO

Sehr beliebt

LASCANA Kurzarmshirt mit Zierknopfleiste aus weicher Piqué-...
ab 29,99 € (29,99 €/1 Stück)
OTTO

Sehr beliebt

THE SET Langarmshirt The Set Rundhals-Langarmshirts, Ser-Pack (...
45,00 €
OTTO

Sehr beliebt

ONLY Shirtbluse ONLVIC S/S AOP TOP NOOS PTM mit Allover Druck
ab 7,90 € UVP 46,99 €
-53%
OTTO

Sehr beliebt

LASCANA Kurzarmshirt mit Biesenstruktur, bügelfreies T-Shirt...
ab 29,99 €
OTTO

Fast ausverkauft

LASCANA Paillettenshirt mit weiten Ärmeln, festliches Glitzershirt,...
ab 49,99 €
OTTO

With interleaving both rankers are presented to the same user simultaneously

A

**VIVANCE T-Shirt mit modischer Streifenstruktur aus weicher...
ab 19,99 € (19,99 €/1 Stk)
OTTO DE**

**VIVANCE 3/4-Arm-Shirt aus bügelfreier Crêpe-Qualität
ab 26,99 € (26,99 €/1 Stk)
OTTO DE**

**VIVANCE T-Shirt mit modischer Streifenstruktur (Packung) aus...
ab 34,98 € (17,49 €/1 Stk)
OTTO DE**

**LASCANA Kurzarmshirt mit Biesenstruktur, bügelfreies T-Shirt...
ab 29,99 €
OTTO DE**

**LASCANA Kurzarmshirt mit Spitzendetail im Rücken, bequemes...
ab 29,99 €
OTTO DE**

**ELBSAND T-Shirt aus weichem Jersey, Kurzarmshirt, sportlich und...
29,99 €
OTTO DE**

B

**RMK T-Shirt Damen Shirt Top Sommer Basic Be Happy Glücklich a...
ab 9,90 € (UVP 29,99 €)
-67%
OTTO DE**

**LASCANA Kurzarmshirt mit Zierknopfleiste aus weicher Piqué-...
ab 29,99 € (29,99 €/1 Stk)
OTTO DE**

**THE SET Langarmshirt The Set Rundhals-Langarmshirts, Ser-Pack (...
45,00 €
OTTO DE**

**ONLY Shirtbluse ONLVIC S/S AOP TOP NOOS PTM mit Allover Druck
ab 7,90 € (UVP 46,99 €)
-53%
OTTO DE**

**LASCANA Kurzarmshirt mit Biesenstruktur, bügelfreies T-Shirt...
ab 29,99 €
OTTO DE**

**LASCANA Paillettenshirt mit weiten Ärmeln, festliches Glitzershirt,...
ab 49,99 €
OTTO DE**

With interleaving both rankers are presented to the same user simultaneously

A

**VIVANCE T-Shirt mit modischer Streifenstruktur aus weicher...
ab 19,99 € (19,99 €/1 Stk)
OTTO DE**

**VIVANCE 3/4-Arm-Shirt aus bügelfreier Crêpe-Qualität
ab 26,99 € (26,99 €/1 Stk)
OTTO DE**

**VIVANCE T-Shirt mit modischer Streifenstruktur (Packung) aus...
ab 26,99 € (17,49 €/1 Stk)
OTTO DE**

**LASCANA Kurzarmshirt mit Biesenstruktur, bügelfreies T-Shirt...
ab 29,99 €
OTTO DE**

**LASCANA Kurzarmshirt mit Spitzendetail im Rücken, bequemes...
ab 29,99 €
OTTO DE**

**ELBSAND T-Shirt aus weichem Jersey, Kurzarmshirt, sportlich und...
29,99 €
OTTO DE**

INTERLEAVED LIST

**VIVANCE T-Shirt mit modischer Streifenstruktur aus weicher...
ab 19,99 € (19,99 €/1 Stk)
OTTO DE**

**RMK T-Shirt Damen Shirt Top Sommer Basic Be Happy Glücklich a...
ab 9,90 € UVP 29,99 € -67%
OTTO DE**

**VIVANCE 3/4-Arm-Shirt aus bügelfreier Crêpe-Qualität
ab 26,99 € (26,99 €/1 Stk)
OTTO DE**

**LASCANA Kurzarmshirt mit Zierknopfleiste aus weicher Piqué-...
ab 29,99 € (29,99 €/1 Stk)
OTTO DE**

**ELBSAND T-Shirt aus weichem Jersey, Kurzarmshirt, sportlich und...
29,99 €
OTTO DE**

**VIVANCE T-Shirt mit modischer Streifenstruktur (Packung) aus...
ab 26,99 € (17,49 €/1 Stk)
OTTO DE**

**THE SET Langarmshirt The Set Rundhals-Langarmshirts, Ser-Pack (...
45,00 €
OTTO DE**

B

**RMK T-Shirt Damen Shirt Top Sommer Basic Be Happy Glücklich a...
ab 9,90 € UVP 29,99 € -67%
OTTO DE**

**LASCANA Kurzarmshirt mit Zierknopfleiste aus weicher Piqué-...
ab 29,99 € (29,99 €/1 Stk)
OTTO DE**

**THE SET Langarmshirt The Set Rundhals-Langarmshirts, Ser-Pack (...
45,00 €
OTTO DE**

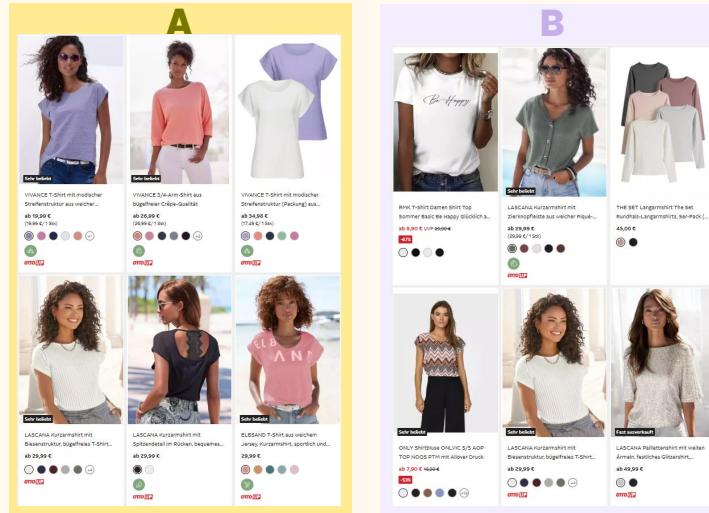
**ONLY Shirtbluse ONLVIC 5/5 AOP TOP NOOS PTM mit Allover Druck
ab 7,90 € 46,99 € -53%
OTTO DE**

**LASCANA Kurzarmshirt mit Biesenstruktur, bügelfreies T-Shirt...
ab 29,99 €
OTTO DE**

**LASCANA Paillettenshirt mit weiten Ärmeln, festliches Glitzershirt,...
ab 49,99 €
OTTO DE**

Interleaving is much more sensitive in detecting effects

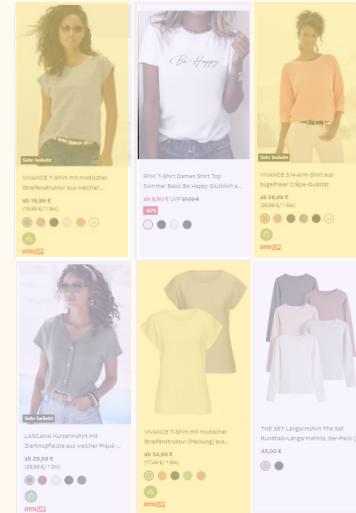
BETWEEN-SUBJECT DESIGN



100% TRAFFIC



WITHIN-SUBJECT DESIGN



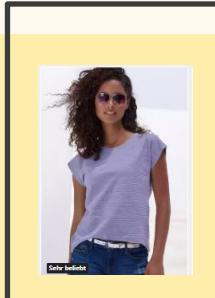
10% TRAFFIC



OTTO

Products are placed on the list following a specific logic

A



B



RESULTING LIST



OTTO

Products are placed on the list following a specific logic

A



B



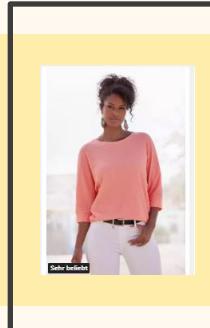
RESULTING LIST



OTTO

Products are placed on the list following a specific logic

A



B



RESULTING LIST



OTTO

Products are placed on the list following a specific logic

A



B



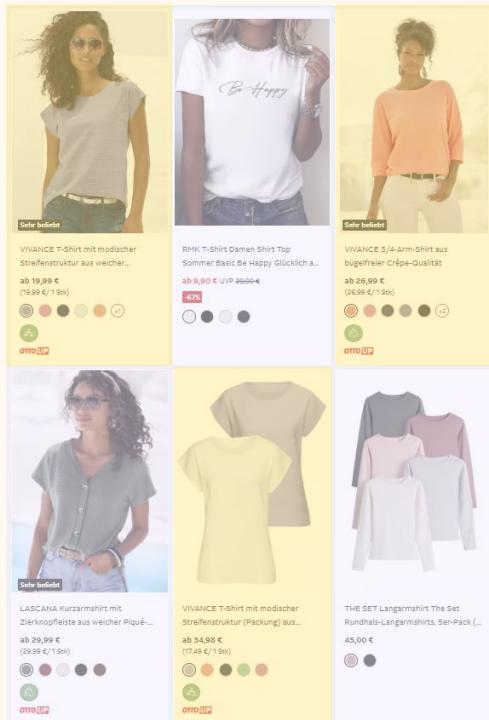
RESULTING LIST



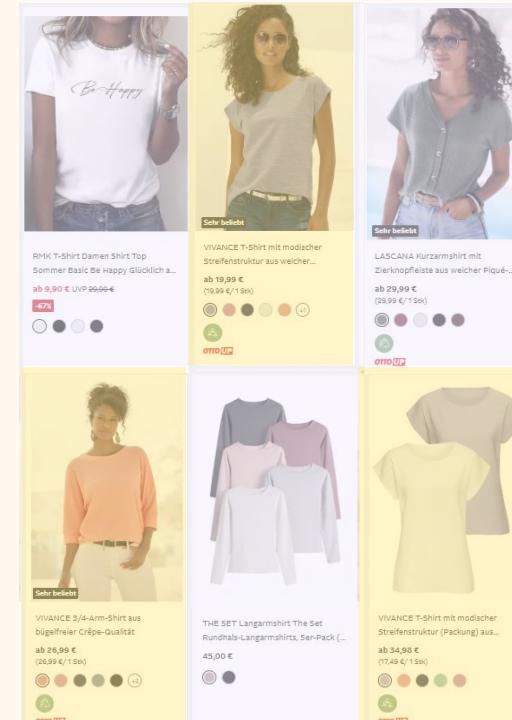
OTTO

Bias #1: The ranker that starts has an advantage

RANKER A STARTING

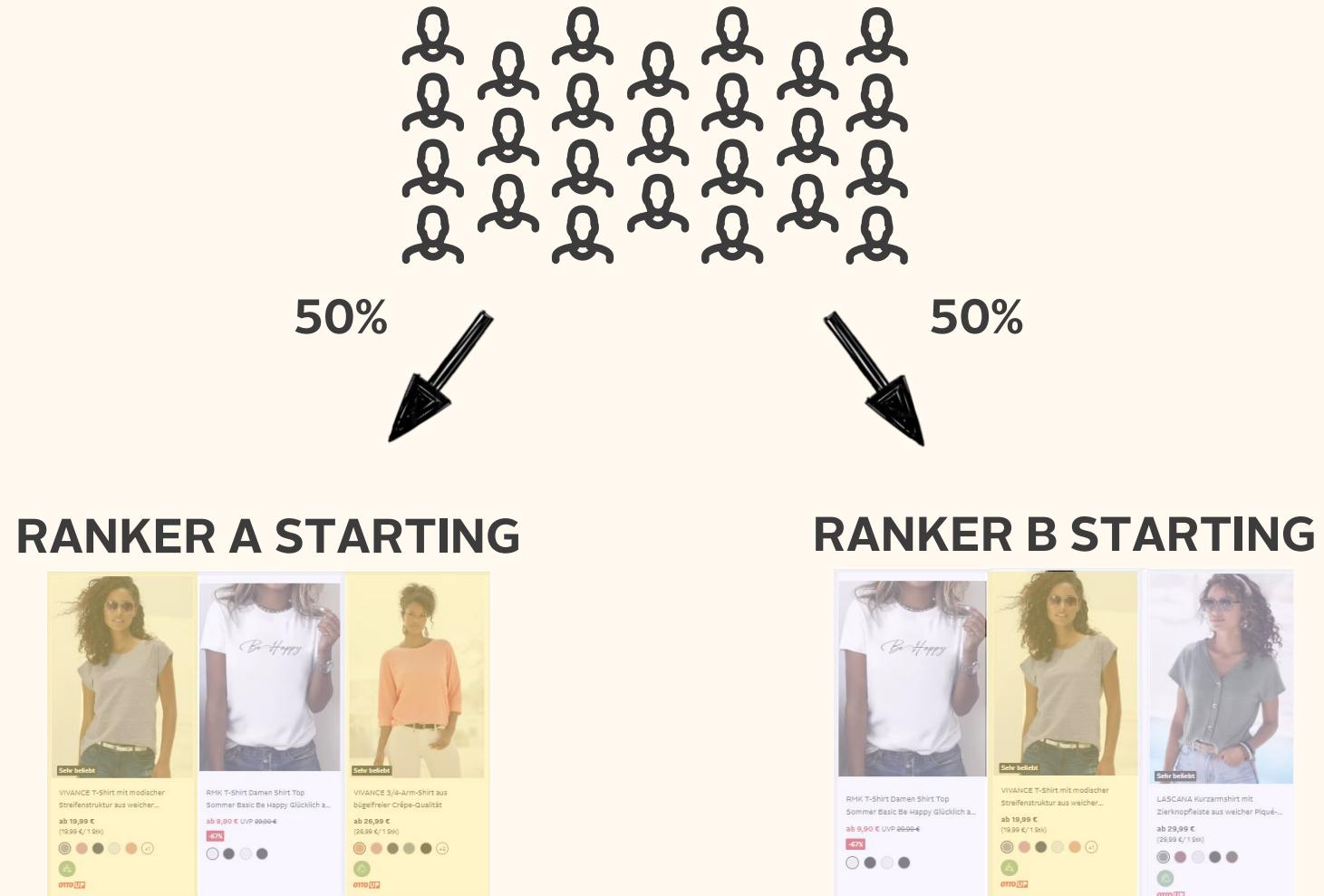


RANKER B STARTING



OTTO

We distribute the advantage evenly among rankers



OTTO

Bias #2: Ranker similarity can lead to an imbalanced list

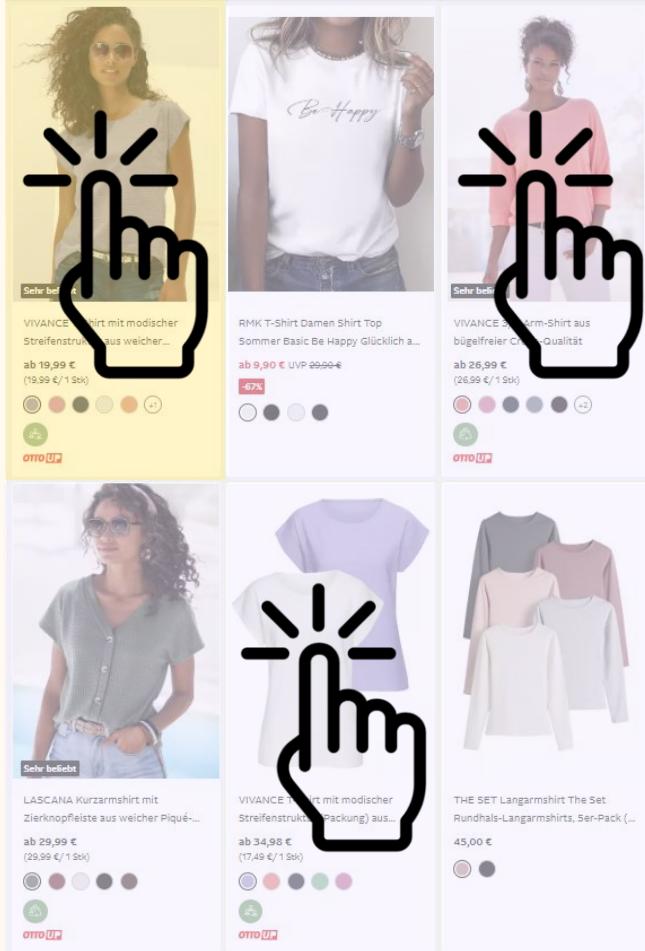


RESULTING LIST



OTTO

In the evaluation of ranker preference we account for imbalanced lists



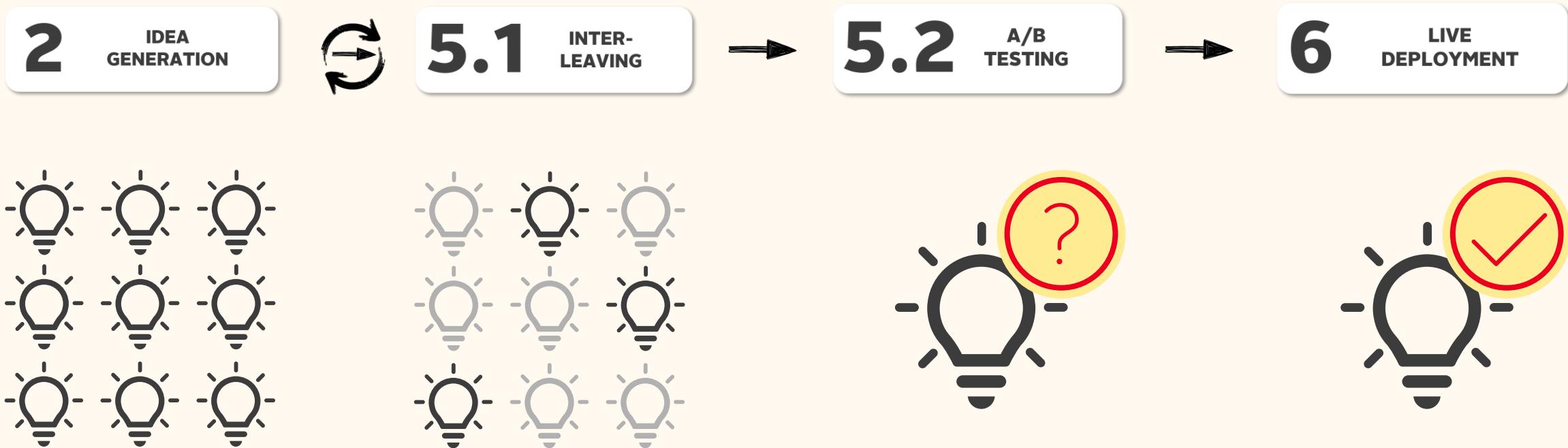
CREDIT = WINS * VISIBILITY RATIO

$$\text{CREDIT RANKER A} = 1 * \frac{6}{1} = 6$$

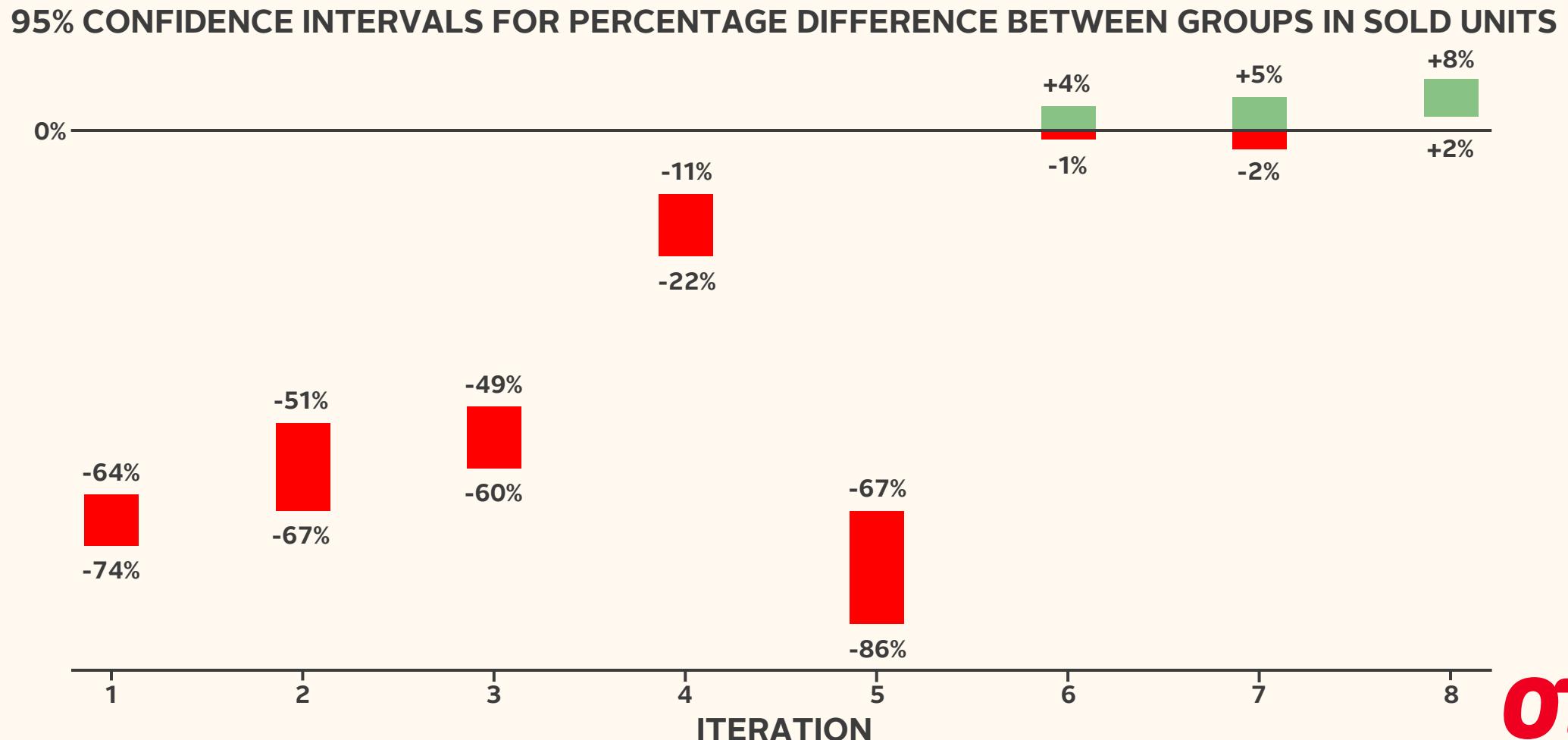
$$\text{CREDIT RANKER B} = 2 * \frac{6}{5} = 2,4$$

OTTO

With a two phase procedure we speed up our iteration cycles



Within 8 weeks we developed a successful deep learning ranking model



We still see potential for improvements

CORRELATION OF INTERLEAVING AND A/B TEST RESULTS

ADD
FOR
LATER

0.68

UNITS

0.47

ORDERS

0.60

OTTO

Hypothesis: using position differences will improve the correlation

A



B



Pos 50

INTERLEAVED LIST



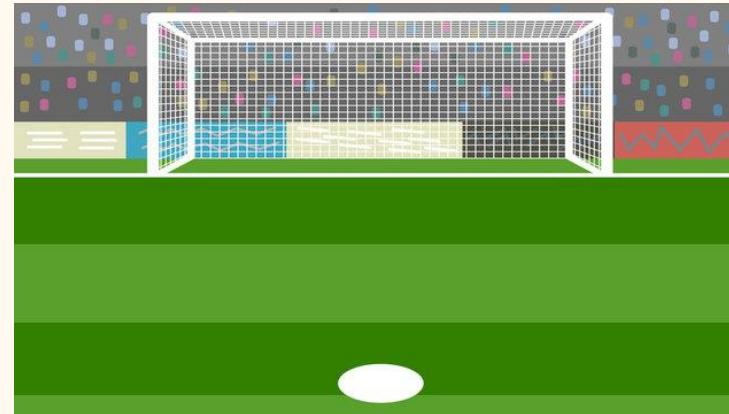
OTTO

Interleaving relates to A/B testing as penalty shootout relates to a football match

A/B TEST



INTERLEAVING

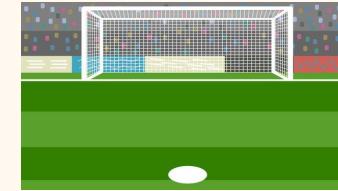


Interleaving relates to A/B testing as penalty shootout relates to a football match

A/B TEST



INTERLEAVING



P
R
O
B
E
L
M
S



The player by itself decides the outcome



The score doesn't represent the „classic“ outcome

S
U
M
M
A
R
Y



More efficient to get results



Classic match is still closer to reality

Big shoutout to Andreas Wagenmann



**Freelance Senior Data Scientist
& ML / Data / Search Engineer**

OTTO



otto