HOCHSCHULE MITTWEIDA
University of Applied Sciences

Cognitive Science and Machine Learning

Mert Saruhan, B.Sc.

November 13, 2023

hs-mittweida.de

Agenda

What is Cognitive Bias?

What is Cognitive Bias?

Introduction

1. Bias created by human cognition
2. Has an active role in decision making
3. Not always logical
4. Notation: $B(q|p)$, How strongly one belives q occurs after observing p
5. $0 \leq B(q|p) \leq 1$

What is Cognitive Bias?

Introduction

1. Bias created by human cognition
2. Has an active role in decision making
3. Not always logical
4. Notation: $B(q|p)$, How strongly one belives q occurs after observing p
5. $0 \leq B(q|p) \leq 1$

HOCHSCHULE MITTWEIDA

What is Cognitive Bias?

Introduction

1. Bias created by human cognition
2. Has an active role in decision making
3. Not always logical
4. Notation: $B(q|p)$, How strongly one belives q occurs after observing p
5. $0 \leq B(q|p) \leq 1$

What is Cognitive Bias?

Introduction

1. Bias created by human cognition
2. Has an active role in decision making
3. Not always logical
4. Notation: $B(q|p)$, How strongly one belives q occurs after observing p
5. $0 \leq B(q|p) \leq 1$

What is Cognitive Bias?

Introduction

1. Bias created by human cognition
2. Has an active role in decision making
3. Not always logical
4. Notation: $B(q|p)$, How strongly one belives q occurs after observing p
5. $0 \leq B(q|p) \leq 1$

What is Cognitive Bias?

Types of cognitive bias we use

- **Symmetry Bias**
  Example: 'If the weather was rainy, then the ground is wet'
  $\Longrightarrow$ 'Only if the ground is wet, then the weather was rainy a while ago' [1]
- **Mutual Exclusitivity Bias**
  Example: 'if you do not clean your room, then you will not be allowed to play'
  $\Longrightarrow$ 'if I clean up my room, then my mom will allow me to play' [2]

What is Cognitive Bias?

Types of cognitive bias we use

- **Symmetry Bias**
  Example: 'If the weather was rainy, then the ground is wet'
  $\Longrightarrow$ 'Only if the ground is wet, then the weather was rainy a while ago' [1]
- **Mutual Exclusitivity Bias**
  Example: 'if you do not clean your room, then you will not be allowed to play'
  $\Longrightarrow$ 'If I clean up my room, then my mom will allow me to play' [2]

What is Cognitive Bias?

Types of cognitive bias we use

- **Symmetry Bias**
  Example: 'If the weather was rainy, then the ground is wet'
  $\Longrightarrow$ 'Only if the ground is wet, then the weather was rainy a while ago' [1]
- **Mutual Exclusitivity Bias**
  Example: 'if you do not clean your room, then you will not be allowed to play'
  $\Longrightarrow$ 'if I clean up my room, then my mom will allow me to play' [2]

What is Cognitive Bias?

Types of cognitive bias we use

- **Symmetry Bias**
  Example: 'If the weather was rainy, then the ground is wet'
  $\Longrightarrow$ 'Only if the ground is wet, then the weather was rainy a while ago' [1]
- **Mutual Exclusitivity Bias**
  Example: 'if you do not clean your room, then you will not be allowed to play'
  $\Longrightarrow$ 'if I clean up my room, then my mom will allow me to play' [2]

What is Cognitive Bias?

Illogical bias

p: ‘The shoe is white’
q: ‘A star is printed on it’
$p \implies q$: ‘If the shoe is white, then a star is printed on it’ [**tan18**]

Symmetry Bias

$q \implies p$: ‘If a star is printed on a shoe, then the shoe is white’ [**tan18**]

Mutual Exclusitivity Bias

$\neg p \implies \neg q$: ‘If the shoe is not white, then a star is not printed on it’ [**tan18**]

What is Cognitive Bias?

Illogical bias

p: 'The shoe is white'
q: 'A star is printed on it'
$p \implies q$: 'If the shoe is white, then a star is printed on it' [**tan18**]

Symmetry Bias

$q \implies p$: 'If a star is printed on a shoe, then the shoe is white' [**tan18**]

Mutual Exclusitivity Bias

$\neg p \implies \neg q$: 'If the shoe is not white, then a star is not printed on it' [**tan18**]

What is Cognitive Bias?

Illogical bias

p: 'The shoe is white'
q: 'A star is printed on it'
$p \implies q$: 'If the shoe is white, then a star is printed on it' [**tan18**]

Symmetry Bias

$q \implies p$: 'If a star is printed on a shoe, then the shoe is white' [**tan18**]

Mutual Exclusivity Bias

$\neg p \implies \neg q$: 'If the shoe is not white, then a star is not printed on it' [**tan18**]

What is Cognitive Bias?

Illogical bias

p: 'The shoe is white'
q: 'A star is printed on it'
$p \implies q$: 'If the shoe is white, then a star is printed on it' [**tan18**]

Symmetry Bias

$q \implies p$: 'If a star is printed on a shoe, then the shoe is white' [**tan18**]

Mutual Exclusitivity Bias

$\neg p \implies \neg q$: 'If the shoe is not white, then a star is not printed on it' [**tan18**]

What is Cognitive Bias?

Properties and biases

- **Symmetry Bias (S)**: $B(q|p) \sim B(p|q)$
- **Mutual Exclusitivity Bias (MX)**: $B(q|p) \sim B(\neg q|\neg p)$
- **The law of excluded middle (XM)**: $B(q|p) \sim 1 - B(\neg q|p)$
- **Estimation relativity (ER)**: $B(q|p) \sim 1 - B(q|\neg p)$

Note: Adapted from [**tak10**]

What is Cognitive Bias?

Properties and biases

- **Symmetry Bias (S)**: $B(q|p) \sim B(p|q)$
- **Mutual Exclusitivity Bias (MX)**: $B(q|p) \sim B(\neg q|\neg p)$
- **The law of excluded middle (XM)**: $B(q|p) \sim 1 - B(\neg q|p)$
- **Estimation relativity (ER)**: $B(q|p) \sim 1 - B(q|\neg p)$

Note: Adapted from [**tak10**]

What is Cognitive Bias?

Properties and biases

- **Symmetry Bias (S):** $B(q|p) \sim B(p|q)$
- **Mutual Exclusitivity Bias (MX):** $B(q|p) \sim B(\neg q|\neg p)$
- **The law of excluded middle (XM):** $B(q|p) \sim 1 - B(\neg q|p)$
- **Estimation relativity (ER):** $B(q|p) \sim 1 - B(q|\neg p)$

Note: Adapted from [**tak10**]

What is Cognitive Bias?

Properties and biases

- **Symmetry Bias (S):** $B(q|p) \sim B(p|q)$
- **Mutual Exclusivity Bias (MX):** $B(q|p) \sim B(\neg q|\neg p)$
- **The law of excluded middle (XM):** $B(q|p) \sim 1 - B(\neg q|p)$
- **Estimation relativity (ER):** $B(q|p) \sim 1 - B(q|\neg p)$

Note: Adapted from [**tak10**]

ML Implementation

ML Implementation

Interpretation

	q	$\neg q$
p	a	b
$\neg p$	c	d

Table: Co-occurence frequency **[man21]**

	$L(x) = L(w^x)$	$L(x) \neq L(w^x)$
$L(w_i) = L(w^x)$	a	b
$L(w_i) \neq L(w^x)$	c	d

Table: ML implementation of the co-occurence frequency table **[man21]**

- x: sample
- w_i: ith prototype
- w_x: winner prototype of sample x
- $L(y)$: label of y

ML Implementation

Interpretation

	q	$\neg q$
p	a	b
$\neg p$	c	d

Table: Co-occurence frequency [**man21**]

	$L(x) = L(w^x)$	$L(x) \neq L(w^x)$
$L(w_i) = L(w^x)$	a	b
$L(w_i) \neq L(w^x)$	c	d

Table: ML implementation of the co-occurence frequency table [**man21**]

- x: sample
- w_i: ith prototype
- w_x: winner prototype of sample x
- $L(y)$: label of y

ML Implementation

Interpretation

	q	$\neg q$
p	a	b
$\neg p$	c	d

Table: Co-occurence frequency [**man21**]

	$L(x) = L(w^x)$	$L(x) \neq L(w^x)$
$L(w_i) = L(w^x)$	a	b
$L(w_i) \neq L(w^x)$	c	d

Table: ML implementation of the co-occurence frequency table [**man21**]

- x: sample
- w_i: ith prototype
- w_x: winner prototype of sample x
- $L(y)$: label of y

ML Implementation

Updating learning rates

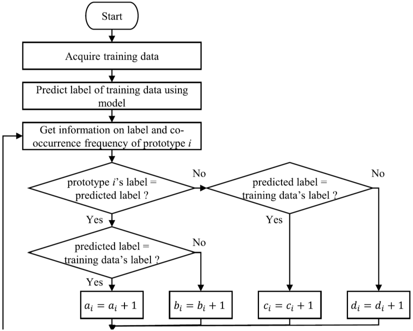


Figure: Learning rate update flowchart part 1 [**tak10**]

ML Implementation

Updating learning rates

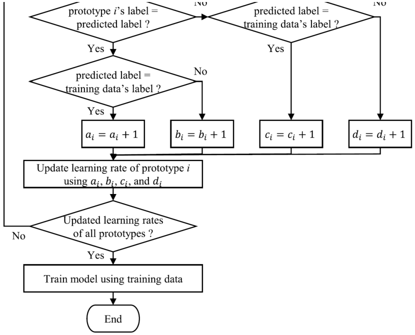


Figure: Learning rate update flowchart part 2 [**tak10**]

ML Implementation

Updating learning rates

- $R_i(a_i, b_i, c_i, d_i, t)$: Causal relationship between events for i^{th} prototype at time t
- $\epsilon_i(t) = 1 - R_i(t)$: Local learning rate of i^{th} prototype at time t
- Each prototype share learning rate with their class!

ML Implementation

Updating learning rates

- $R_i(a_i, b_i, c_i, d_i, t)$: Causal relationship between events for i^{th} prototype at time t
- $\epsilon_i(t) = 1 - R_i(t)$: Local learning rate of i^{th} prototype at time t
- Each prototype share learning rate with their class!

ML Implementation

Updating learning rates

- $R_i(a_i, b_i, c_i, d_i, t)$: Causal relationship between events for i^{th} prototype at time t
- $\epsilon_i(t) = 1 - R_i(t)$: Local learning rate of i^{th} prototype at time t
- Each prototype share learning rate with their class!

Optimizers

Optimizers

Loose Symmetry (LS)

- $R_i^{\mathrm{LS}}(t) = \dfrac{a_i(t) + \frac{b_i(t)}{b_i(t)+d_i(t)} d_i(t)}{a_i(t) + \frac{b_i(t)}{b_i(t)+d_i(t)} d_i(t) + b_i(t) + \frac{a_i(t)}{a_i(t)+c_i(t)} c_i(t)}$ [**tak10**, **man21**]
- Satisfies XM and loosely satisfies S, MX, and ER [**tak10**]
- Has better results than other cognitive bias optimizers [3]

HOCHSCHULE MITTWEIDA

Optimizers

Loose Symmetry (LS)

- $R_i^{\mathrm{LS}}(t) = \frac{a_i(t)+\frac{b_i(t)}{b_i(t)+d_i(t)}d_i(t)}{a_i(t)+\frac{b_i(t)}{b_i(t)+d_i(t)}d_i(t)+b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ [**tak10**, **man21**]
- Satisfies XM and loosely satisfies S, MX, and ER [**tak10**]
- Has better results than other cognitive bias optimizers [3]

Optimizers

Loose Symmetry (LS)

- $R_i^{\mathrm{LS}}(t) = \dfrac{a_i(t) + \frac{b_i(t)}{b_i(t)+d_i(t)} d_i(t)}{a_i(t) + \frac{b_i(t)}{b_i(t)+d_i(t)} d_i(t) + b_i(t) + \frac{a_i(t)}{a_i(t)+c_i(t)} c_i(t)}$ [**tak10**, **man21**]
- Satisfies XM and loosely satisfies S, MX, and ER [**tak10**]
- Has better results than other cognitive bias optimizers [3]

Optimizers

Loose Symmetry under Rarity (LSR)

- Assumption: The events p and q are small, hence the correlation of any two events is unlikely, $d(t) \to \infty$ [2]
- Example: The correlation between any random event and you starting you car in the morning [2]
- $R_i^{\mathrm{LSR}}(t) = \lim_{d_i(t)\to\infty} R_i^{\mathrm{LS}}(t)$
- $R_i^{\mathrm{LS}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ **[tak10]**
- Satisfies ? [3]

Optimizers

Loose Symmetry under Rarity (LSR)

- Assumption: The events p and q are small, hence the correlation of any two events is unlikely, $d(t) \to \infty$ [2]
- Example: The correlation between any random event and you starting you car in the morning [2]
- $R_i^{\mathrm{LSR}}(t) = \lim_{d_i(t)\to\infty} R_i^{\mathrm{LS}}(t)$
- $R_i^{\mathrm{LS}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ [tak10]
- Satisfies ? [3]

Optimizers

Loose Symmetry under Rarity (LSR)

- Assumption: The events p and q are small, hence the correlation of any two events is unlikely, $d(t) \to \infty$ [2]
- Example: The correlation between any random event and you starting you car in the morning [2]
- $R_i^{\text{LSR}}(t) = \lim_{d_i(t)\to\infty} R_i^{\text{LS}}(t)$
- $R_i^{\text{LS}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ **[tak10]**
- Satisfies ? [3]

HOCHSCHULE MITTWEIDA

Optimizers

Loose Symmetry under Rarity (LSR)

- Assumption: The events p and q are small, hence the correlation of any two events is unlikely, $d(t) \to \infty$ [2]
- Example: The correlation between any random event and you starting you car in the morning [2]
- $R_i^{\text{LSR}}(t) = \lim_{d_i(t) \to \infty} R_i^{\text{LS}}(t)$
- $R_i^{\text{LS}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ **[tak10]**
- Satisfies ? [3]

Optimizers

Loose Symmetry under Rarity (LSR)

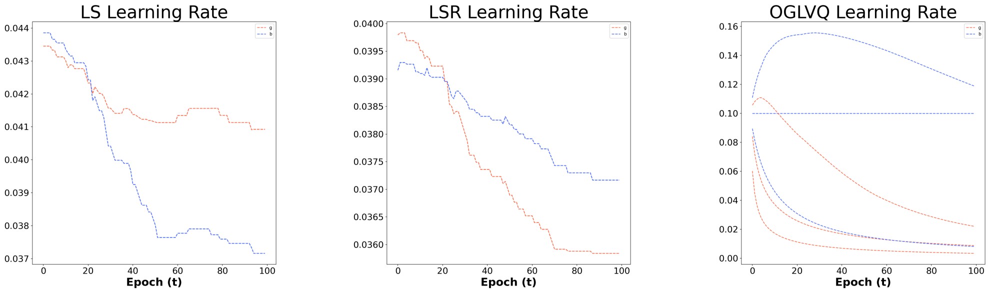
- Assumption: The events p and q are small, hence the correlation of any two events is unlikely, $d(t) \to \infty$ [2]
- Example: The correlation between any random event and you starting you car in the morning [2]
- $R_i^{\text{LSR}}(t) = \lim_{d_i(t) \to \infty} R_i^{\text{LS}}(t)$
- $R_i^{\text{LS}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ **[tak10]**
- Satisfies ? [3]

Results

Results

Balanced dataset

Dataset: Ionosphere dataset [4]



Be careful with y-axis, since they do not share common size

Conclusion

Bibliography I

[1] S. Shinohara, R. Taguchi, K. Katsurada, and T. Nitta, "A model of belief formation based on causality and application to n-armed bandit problem (in japanese)," *Transactions of the Japanese Society for Artificial Intelligence*, vol. 22, no. 1, pp. 58–68, 2007. DOI: `10.1527/tjsai.22.58`.

[2] M. Hattori and M. Oaksford, "Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis," *Cognitive Science: A Multidisciplinary Journal*, vol. 31, no. 5, pp. 765–814, 2007. DOI: `10.1080/03640210701530755`.

[3] smt, "Smt," *Scientific Reports*, vol. 11, no. 1, 2021.

[4] ion, "Ion," *Scientific Reports*, vol. 11, no. 1, 2021.

Thank You

Mert Saruhan, B.Sc.

msaruhan@hs-mittweida.de

Mathematics for Network and Data Science (MA20w1-M)

HOCHSCHULE MITTWEIDA
University of Applied Sciences

Hochschule Mittweida
University of Applied Sciences
Technikumplatz 17 | 09648 Mittweida

hs-mittweida.de