HOCHSCHULE
MITTWEIDA
University of Applied Sciences

Cognitive Science and Machine Learning

Mert Saruhan, B.Sc.

December 6, 2023

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

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

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

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

What is Cognitive Bias?

Types of cognitive bias we use

Symmetry Bias

Example: ‘If the weather was rainy, then the ground is wet’
$(p \implies q)$
$\implies$ ‘Only if the ground is wet, then the weather was rainy a while ago’ [1]
$(q \implies p)$

What is Cognitive Bias?

Types of cognitive bias we use

Symmetry Bias

Example: 'If the weather was rainy, then the ground is wet'

$(p \implies q)$

$\implies$ 'Only if the ground is wet, then the weather was rainy a while ago' [1]

$(q \implies p)$

What is Cognitive Bias?

Types of cognitive bias we use

Mutual Exclusivity Bias

Example: 'If you do not clean your room, then you will not be allowed to play' $(p \implies q)$

$\implies$ 'If I clean up my room, then my mom will allow me to play' [2] $(\neg p \implies \neg q)$

What is Cognitive Bias?

Types of cognitive bias we use

Mutual Exclusitivity Bias
Example: 'If you do not clean your room, then you will not be allowed to play'
$(p \implies q)$
$\implies$ 'If I clean up my room, then my mom will allow me to play' [2]
$(\neg p \implies \neg q)$

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' [3]

Symmetry Bias

$q \implies p$: 'If a star is printed on a shoe, then the shoe is white' [3]

Mutual Exclusivity Bias

$\neg p \implies \neg q$: 'If the shoe is not white, then a star is not printed on it' [3]

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’ [3]

Symmetry Bias

$q \implies p$: ‘If a star is printed on a shoe, then the shoe is white’ [3]

Mutual Exclusitivity Bias

$\neg p \implies \neg q$: ‘If the shoe is not white, then a star is not printed on it’ [3]

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' [3]

Symmetry Bias

$q \implies p$: 'If a star is printed on a shoe, then the shoe is white' [3]

Mutual Exclusitivity Bias

$\neg p \implies \neg q$: 'If the shoe is not white, then a star is not printed on it' [3]

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' [3]

Symmetry Bias
$q \implies p$: 'If a star is printed on a shoe, then the shoe is white' [3]

Mutual Exclusitivity Bias
$\neg p \implies \neg q$: 'If the shoe is not white, then a star is not printed on it' [3]

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 [4]

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 [4]

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 [4]

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 [4]

ML Implementation

ML Implementation

Interpretation

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

Table: Co-occurence frequency [5]

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

Table: ML implementation of the co-occurence frequency table [5]

- x: sample
- w_i: i^{th} 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 [5]

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

Table: ML implementation of the co-occurence frequency table [5]

- x: sample
- w_i: i^{th} 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 [5]

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

Table: ML implementation of the co-occurence frequency table [5]

- x: sample
- w_i: i^{th} prototype
- w_x: winner prototype of sample x
- $L(y)$: label of y

ML Implementation

Updating learning rates

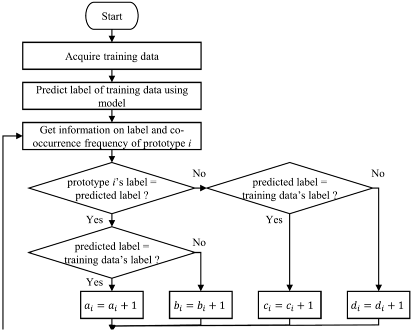


Figure: Learning rate update flowchart part 1 [4]

ML Implementation

Updating learning rates

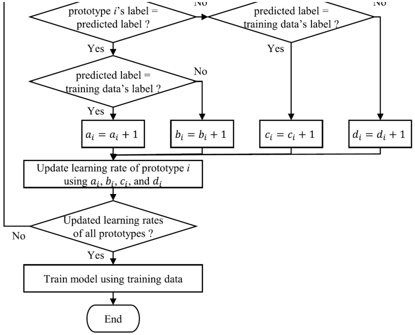


Figure: Learning rate update flowchart part 2 [4]

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)

- $R_i^{\mathrm{LS}}(t) = \dfrac{a_i(t)+\frac{b_i(t)}{b_i(t)+d_i(t)}d_i(t)}{a_i(t)+\frac{b_i(t)}{b_i(t)+d_i(t)}d_i(t)+b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ [4, 5]
- Satisfies XM and loosely satisfies S, MX, and ER [4]
- Has better results than other cognitive bias optimizers in tested datasets [4]

Optimizers

Loose Symmetry (LS)

- $R_i^{\mathrm{LS}}(t) = \dfrac{a_i(t) + \frac{b_i(t)}{b_i(t)+d_i(t)} d_i(t)}{a_i(t) + \frac{b_i(t)}{b_i(t)+d_i(t)} d_i(t) + b_i(t) + \frac{a_i(t)}{a_i(t)+c_i(t)} c_i(t)}$ [4, 5]
- Satisfies XM and loosely satisfies S, MX, and ER [4]
- Has better results than other cognitive bias optimizers in tested datasets [4]

Optimizers

Loose Symmetry (LS)

- $R_i^{\mathrm{LS}}(t) = \dfrac{a_i(t) + \frac{b_i(t)}{b_i(t)+d_i(t)} d_i(t)}{a_i(t) + \frac{b_i(t)}{b_i(t)+d_i(t)} d_i(t) + b_i(t) + \frac{a_i(t)}{a_i(t)+c_i(t)} c_i(t)}$ [4, 5]
- Satisfies XM and loosely satisfies S, MX, and ER [4]
- Has better results than other cognitive bias optimizers in tested datasets [4]

Optimizers

Loose Symmetry under Rarity (LSR)

- Assumption: The events p and q are small, hence the correlation of any two events is unlikely, $d(t) \rightarrow \infty$ [2]
- Example: The correlation between any random event and you starting you car in the morning [2]
- $R_i^{\text{LSR}}(t) = \lim_{d_i(t)\rightarrow\infty} R_i^{\text{LS}}(t)$
- $R_i^{\text{LSR}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ [4]
- Satisfies S [4]

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]
- $R_i^{\mathrm{LSR}}(t) = \lim_{d_i(t)\to\infty} R_i^{\mathrm{LS}}(t)$
- $R_i^{\mathrm{LSR}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ [4]
- Satisfies S [4]

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]
- $R_i^{\mathrm{LSR}}(t) = \lim_{d_i(t)\to\infty} R_i^{\mathrm{LS}}(t)$
- $R_i^{\mathrm{LSR}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ [4]
- Satisfies S [4]

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]
- $R_i^{\text{LSR}}(t) = \lim_{d_i(t) \to \infty} R_i^{\text{LS}}(t)$
- $R_i^{\text{LSR}}(t) = \dfrac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ [4]
- Satisfies S [4]

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]
- $R_i^{\text{LSR}}(t) = \lim_{d_i(t) \to \infty} R_i^{\text{LS}}(t)$
- $R_i^{\text{LSR}}(t) = \frac{a_i(t)+b_i(t)}{a_i(t)+2b_i(t)+\frac{a_i(t)}{a_i(t)+c_i(t)}c_i(t)}$ [4]
- Satisfies S [4]

Results

Results

- Compare the performance cognitive models with OGVLQ model
- OGLVQ model stands for optimized GLVQ [6]
- For balanced dataset, use accuracy as score
- For Imbalanced dataset, use F1 score as score
- Prototype selection is optimized for OGLVQ

Results

- Compare the performance cognitive models with OGVLQ model
- OGLVQ model stands for optimized GLVQ [6]
- For balanced dataset, use accuracy as score
- For Imbalanced dataset, use F1 score as score
- Prototype selection is optimized for OGLVQ

Results

- Compare the performance cognitive models with OGVLQ model
- OGLVQ model stands for optimized GLVQ [6]
- For balanced dataset, use accuracy as score
- For Imbalanced dataset, use F1 score as score
- Prototype selection is optimized for OGLVQ

Results

- Compare the performance cognitive models with OGVLQ model
- OGLVQ model stands for optimized GLVQ [6]
- For balanced dataset, use accuracy as score
- For Imbalanced dataset, use F1 score as score
- Prototype selection is optimized for OGLVQ

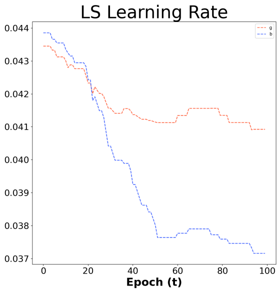
Results

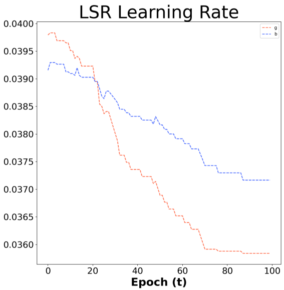
- Compare the performance cognitive models with OGVLQ model
- OGLVQ model stands for optimized GLVQ [6]
- For balanced dataset, use accuracy as score
- For Imbalanced dataset, use F1 score as score
- Prototype selection is optimized for OGLVQ

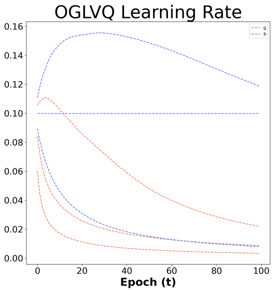
Results

Balanced dataset

Dataset: Ionosphere dataset [7]







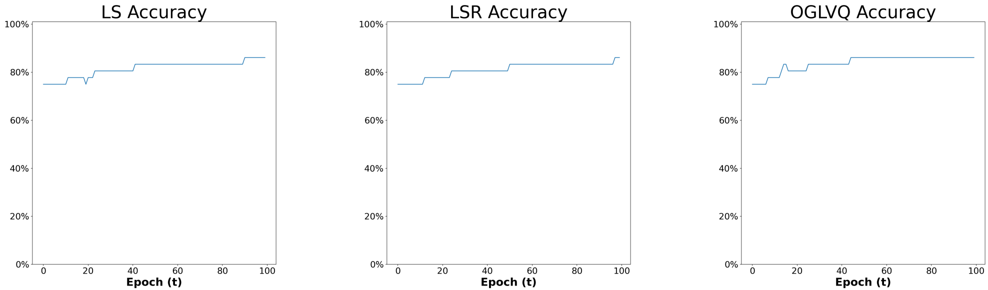
Be careful with y-axis, since they do not share common size

HOCHSCHULE MITTWEIDA

Results

Balanced dataset

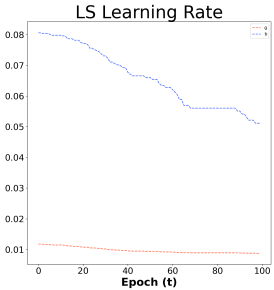
Dataset: Ionosphere dataset [7]

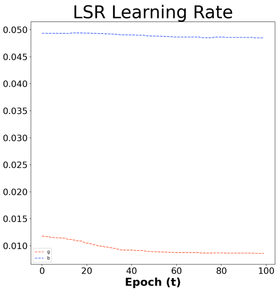


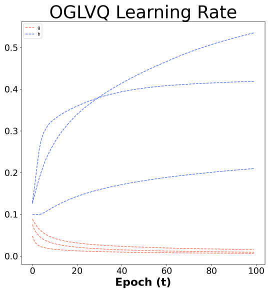
Results

Imbalanced dataset

Dataset: Ionosphere dataset [7]





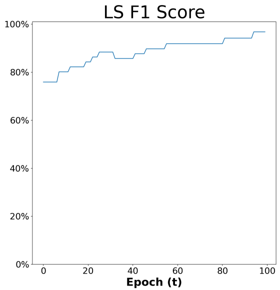


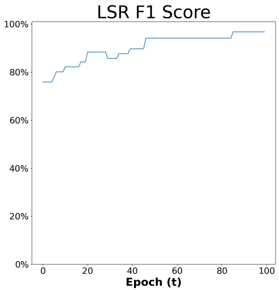
Be careful with y-axis, since they do not share common size

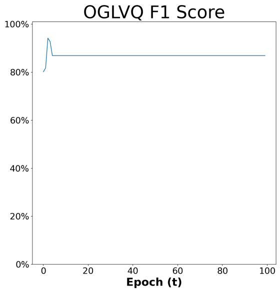
Results

Imbalanced dataset

Dataset: Ionosphere dataset [7]







Conclusion

Cognitive bias optimizer . . .

- use same learning rate for each prototype in the same class
- have better learning rate graphs than OGLVQ
- have better learning performance than OGLVQ

Conclusion

Cognitive bias optimizer . . .

- use same learning rate for each prototype in the same class
- have better learning rate graphs than OGLVQ
- have better learning performance than OGLVQ

Conclusion

Cognitive bias optimizer . . .

- use same learning rate for each prototype in the same class
- have better learning rate graphs than OGLVQ
- have better learning performance than OGLVQ

Further

Need further testing on . . .

- optimized datasets
- noiser datasets

Further

Need further testing on ...

- optimized datasets
- noiser datasets

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] H. Taniguchi, H. Sato, and T. Shirakawa, "A machine learning model with human cognitive biases capable of learning from small and biased

Bibliography II

datasets," *Scientific Reports*, vol. 8, no. 1, 2018. DOI: 10.1038/s41598-018-25679-z.

[4] T. Takahashi, M. Nakano, and S. Shinohara, "Cognitive symmetry: Illogical but rational biases," *Symmetry Culture and Science*, vol. 21, no. 1-3, pp. 275–294, 2010. [Online]. Available: https://www.researchgate.net/publication/285850238_Cognitive_Symmetry_Illogical_but_Rational_Biases.

[5] N. Manome, S. Shinohara, T. Takahashi, Y. Chen, and U. Chung, "Selfincremental learning vector quantization with human cognitive biases," *Scientific Reports*, vol. 11, no. 1, 2021. DOI: 10.1038/s41598-021-83182-4.

Bibliography III

[6] T. Kohonen, *Self-Organizing Maps*. Berlin, Heidelberg: Springer-Verlag, 1995, pp. 175–189.

[7] V. G. Sigillito, S. P. Wing, L. V. Hutton, and K. B. Baker, “Classification of radar returns from the ionosphere using neural networks,” *Johns Hopkins APL Technical Digest*, vol. 10, no. 1, pp. 262–266, 1989.

Thank You

Mert Saruhan, B.Sc.

msaruhan@hs-mittweida.de

Mathematics for Network and Data Science (MA20w1-M)

Hochschule Mittweida
University of Applied Sciences
Technikumplatz 17 | 09648 Mittweida

HOCHSCHULE MITTWEIDA
University of Applied Sciences

hs-mittweida.de