به نام ایزد یکتا

تمرین دوم درس روش پژوهش و ارائه

استاد: دکتر رضا صفابخش

تهیه کننده: بردیا اردکانیان

برسی و مهار مثال‌های خصمانه در یادگیری ماشین

الف)

موضوع پژوهش: برسی و مهار مثال‌های خصمانه در یادگیری ماشین

EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES IN MACHINE LEARNING

واژه‌های کلیدی:

1. Machine Learning
2. Adversarial
3. Adversarial Attacks
4. Adversarial Examples
5. Explaining
6. Harnessing

کلمات کلیدی ذکر شده را در موتورهای مختلف جست و جو می‌کنیم.

1) جست و جو با Adversarial + Machine Learning

الف) جست و جو در Google (5 صفحه اول)

G01: Adversarial machine learning

G02: Wild patterns: Ten years after the rise of adversarial machine learning

G03: Adversarial Machine Learning

G04: Machine Learning in Adversarial Settings

G05: Machine learning in adversarial environments

G06: Machine learning uncertainties with adversarial neural networks

G07: Adversarial Machine Learning Beyond the Image Domain

ب) جست و جو در Yahoo (2 صفحه اول)

Y01: A Taxonomy and Terminology of 21 Adversarial Machine Learning

Y02: Adversarial Machine Learning Attacks and Defense Methods in the Cyber Security Domain

Y03: Adversarial machine learning for cybersecurity and computer vision: Current developments and challenges

پ) جست و جو در Bing (3 صفحه اول)

BO1: Adversarial machine learning (Duplicate)

منابع پیدا شده توسط این موتور جست و جو تکراری و یا نا مرتبط بوده‌اند.

2) جست و جو با Adversarial Examples + Machine Learning:

الف) جست و جو در Google (5 صفحه اول)

G08: Explaining and Harnessing Adversarial Examples

G09: Adversarial Examples: Attacks and Defenses for Deep Learning

G10: Adversarial Examples Are Not Easily Detected: Bypassing Ten Detection Methods

G11: Efficient Adversarial Training with Transferable Adversarial Examples

G12: Understanding Adversarial Examples From the Mutual Influence of Images and Perturbations

G13: Generating Natural Adversarial Examples

G14: Adversarial Examples in Modern Machine Learning: A Review

ب)

شماره	کد مقاله	سال چاپ	ناشر	تعداد ارجاعات مقاله	شاخص h	ضریب تاثیر	اعتبار نویسندگان
1	G01	2011	ACM	445	163	SJR=2.079	12, 59, -, 34, -
2	G02	2018	ACM	30	163	SJR=2.079	38, 73
3	G03	2018	Morgan & Claypool	142	24	SJR=0.927	38, 54
4	G04	2016	IEEE	61	76	SJR=0.53	79, 35, 17
5	G05	2010	Springer Nature	51	17	SJR=0.53	35, 53
6	G06	2019	Springer Nature	31	17	SJR=0.53	10, -, 31, 66
7	G07	2019	IEEE	0	-	-	5, 36, 20
8	G08	2015	-	10630	-	-	79, 53, 31
9	G09	2019	IEEE	395	212	SJR=2.882	10, 12, 8, 81
10	G10	2017	ACM	1297	9	SJR=2.95	31, 95
11	G11	2020	IEEE	39	406	SJR=4.658	5, 2, 8, 75
12	G12	2020	IEEE	49	-	-	12, 11, 5
13	G13	2017	-	399	-	-	6, 8, 40
14	G14	2019	-	44	-	-	3, 18, 4, 15
15	Y01	2019	NIST	38	59	SJR=0.202	12
16	Y02	2021	ACM	13	163	SJR=2.079	10, 41, 67
17	Y03	2020	Wiley Interdisciplinary	6	38	SJR=0.693	17
18	B01	-	-	-	-	-	-

ج)

معیارهایی همچون سال چاپ، اعتبار نویسندگان، اعتبار ناشر، ارجاعات معتبر و ارتباط مستقیم موضوعات با موضوع پژوهشی را می‌توان اضافه کرد. در انتخاب شخصی اعتبار نویسندگاه را اضافه کرده‌ام.

د)

مقالات منتخب با اولویت به شرح ذیل می‌باشند:

1. G08
2. G01
3. G09
4. G11
5. Y02
6. Y01