Mastane Achab, PhD

Abu Dhabi, UAE mastane.achab@gmail.com

Investigating regularization of self-play language models Deep reinforcement learning algorithms for hybrid V2X communication A risk-averse framework for non-stationary stochastic bandits ICDM WS, 12/202 Theory and algorithm for sum-log-concave optimization One-step distributional reinforcement learning A nested matrix-tensor model for noisy multi-view clustering Distributional deep Q-learning with CVaR regression Checkered Regression Robustness via distributional dynamic programming Weighted empirical risk minimization Dimensionality reduction and (bucket) ranking ALT, 03/202 Ranking data with continuous labels Max k-armed bandits on ExtremeHunter and beyond CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem Telécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' ICDM WS, 12/202 ICDM W	Lightlike pseudo-Riemannian adaptive gradient method	preprint, 04/2025
Deep reinforcement learning algorithms for hybrid V2X communication A risk-averse framework for non-stationary stochastic bandits TCDM WS, 12/202 Theory and algorithm for sum-log-concave optimization One-step distributional reinforcement learning A nested matrix-tensor model for noisy multi-view clustering Distributional deep Q-learning with CVaR regression Checkered Regression Checkered Regression Checkered Regression Checkered Regression Checkered Rejevation and dynamic programming Weighted empirical risk minimization Dimensionality reduction and (bucket) ranking Profitable bandits Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postedor researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 2012 - 202 2012 - 202 2012 - 202 2012 - 202 2012 - 202	A Bregman firmly nonexpansive proximal operator for baryconvex optimiza	tion preprint, 11/2024
A risk-averse framework for non-stationary stochastic bandits Theory and algorithm for sum-log-concave optimization One-step distributional reinforcement learning A nested matrix-tensor model for noisy multi-view clustering Distributional deep Q-learning with CVaR regression Checkered Regression Robustness via distributional dynamic programming Weighted empirical risk minimization Dimensionality reduction and (bucket) ranking Profitable bandits ACML, 11/202 Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 ECOLE - 2012 - 203 ECOLE - 2013 - 203 ECOLE - 2015 - 203 ECOLE - 2016 - 203 ECOLE - 2016 - 203 ECOLE - 204 ECOLE - 205 ECO	Investigating regularization of self-play language models	preprint, $04/2024$
Theory and algorithm for sum-log-concave optimization One-step distributional reinforcement learning A nested matrix-tensor model for noisy multi-view clustering Distributional deep Q-learning with CVaR regression Checkered Regression Robustness via distributional dynamic programming Weighted empirical risk minimization Dimensionality reduction and (bucket) ranking ALT, 03/20 Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 Evel Comparis Cachan, France Bécole polytechnique, Palaiseau, France 2012 - 202 Evel Comparis Cachan, France 2015 - 205 Evel Cole polytechnique, Palaiseau, France 2016 - 205 Evel Cole polytechnique, Palaiseau, France 2017 - 205 Evel Cole polytechnique, Palaiseau, France 2018 - 205 Evel Cole polytechnique, Palaiseau, France 2019 - 205 Evel Cole polytechnique, Palaiseau, France 2010 - 205 Evel Cole polytechnique, Palaiseau, France 2011 - 205 Evel Cole polytechnique, Palaiseau, France 2012 - 205 Evel Cole polytechnique, Palaiseau, France 2013 - 205 Evel Cole polytechnique, Palaiseau, France 2016 - 205 Evel Cole polytechnique, Palaiseau, France 2017 - 205 Evel Cole polytechnique, Palaiseau, France 2018 - 205 Evel Cole polytechnique, Palaiseau, France 2019 - 205 Evel Cole polytechnique, Palaiseau, France 2019 - 205 Evel Cole polytechnique, Palaiseau, France 2019 - 205 Evel Cole polytechnique particular dev	Deep reinforcement learning algorithms for hybrid V2X communication	IEEE ICC 2024, 06/2024
One-step distributional reinforcement learning A nested matrix-tensor model for noisy multi-view clustering Distributional deep Q-learning with CVaR regression Checkered Regression Robustness via distributional dynamic programming Weighted empirical risk minimization Dimensionality reduction and (bucket) ranking Profitable bandits Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 2012 - 202 2012 - 203 2013 - 203 2015 - 203 2016 - 203 2016 - 203 2016 - 203 2017 - 203 2018 - 203 2019 - 203 201	A risk-averse framework for non-stationary stochastic bandits	ICDM WS, 12/2023
A nested matrix-tensor model for noisy multi-view clustering Distributional deep Q-learning with CVaR regression Checkered Regression Robustness via distributional dynamic programming Weighted empirical risk minimization Dimensionality reduction and (bucket) ranking Profitable bandits Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 205 2012 - 205 Profitable dempirical risk minimization Preprint, 02/202 Rounding preprint, 02/202 Rounding preprint, 02/202 Rounding preprint, 02/202 REMANN, 10/202 ACML, 11/203 ACML, 11	Theory and algorithm for sum-log-concave optimization	preprint, 09/2023
Distributional deep Q-learning with CVaR regression Checkered Regression Robustness via distributional dynamic programming Weighted empirical risk minimization Dimensionality reduction and (bucket) ranking Profitable bandits Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postedor researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 LT2 / 202 LT2 / 203 LT2 / 203 LT3 / 203 LT3 / 203 LT4 / 203 LT5 / 203 L	One-step distributional reinforcement learning	TMLR, 08/2023
Checkered Regression preprint, 02/202 Robustness via distributional dynamic programming preprint, 12/202 Weighted empirical risk minimization ESANN, 10/202 Dimensionality reduction and (bucket) ranking ALT, 03/203 Profitable bandits ACML, 11/203 Ranking data with continuous labels NeurIPS, 12/203 Max k-armed bandit: on ExtremeHunter and beyond ECML-PKDD, 09/203 CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain 2020 - 202 Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2015 - 203 Ecole polytechnique Palaiseau, France 2015 - 203 Ecole	A nested matrix-tensor model for noisy multi-view clustering	preprint, $05/2023$
Robustness via distributional dynamic programming preprint, 12/202 Weighted empirical risk minimization ESANN, 10/202 Dimensionality reduction and (bucket) ranking ALT, 03/202 Profitable bandits ACML, 11/202 Ranking data with continuous labels ACML, 11/202 Max k-armed bandit: on ExtremeHunter and beyond ECML-PKDD, 09/202 CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain 2020 - 202 Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK 2019 (5 months PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France 2016 - 202 PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France 2015 - 202 MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202	Distributional deep Q-learning with CVaR regression	NeurIPS WS, 12/2022
Weighted empirical risk minimization Dimensionality reduction and (bucket) ranking ALT, 03/20: Profitable bandits ACML, 11/20: Ranking data with continuous labels NeurIPS, 12/20: Max k-armed bandit: on ExtremeHunter and beyond ECML-PKDD, 09/20: EPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 Ecole polytechnique, Palaiseau, France 2015 - 203 Ecole polytechnique, Palaiseau, France 2015 - 203 Ecole polytechnique, Palaiseau, France	Checkered Regression	preprint, $02/2022$
Dimensionality reduction and (bucket) ranking Profitable bandits ACML, 11/202 Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond ECML-PKDD, 09/203 CPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 Ecole polytechnique, Palaiseau, France 2012 - 203 ACML, 11/203 ACML, 11/203 NeurIPS, 12/203 Remarked bandit: on ExtremeHunter and beyond EcML-PKDD, 09/203 ECML-PKDD, 09/203	Robustness via distributional dynamic programming	preprint, 12/2022
Profitable bandits Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond ECML-PKDD, 09/202 REPRIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 Ecole polytechnique, Palaiseau, France 2012 - 203 2012 - 203 2020 - 202 2020 - 202 2019 (5 months	Weighted empirical risk minimization	ESANN, 10/2020
Ranking data with continuous labels Max k-armed bandit: on ExtremeHunter and beyond ECML-PKDD, 09/203 EPERIENCE Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem OUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 203 ECML-PKDD, 09/203 ECML-PKDD, 09/203 ECML-PKDD, 09/203 ECML-PKDD, 09/203 2022 - 202 2020 - 202 2020 - 202 2019 (5 months) 2015 (5 months) Ecole normale supérieure de Cachan, France 2016 - 203 MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 203	Dimensionality reduction and (bucket) ranking	ALT, $03/2019$
Max k-armed bandit: on ExtremeHunter and beyond ECML-PKDD, 09/2020 Technology Innovation Institute (TII), Abu Dhabi, UAE Al Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 2012 - 202 2020 - 202 2019 (5 months 2015 (5 months 2015 - 202 2016 - 202 2016 - 202 2017 - 202 2018 - 202 2019 - 202 2019 (5 months 2010 - 202 2010 - 202 2011 - 202 2011 - 202 2012 - 202 2012 - 202	Profitable bandits	ACML, 11/201
Technology Innovation Institute (TII), Abu Dhabi, UAE Al Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 202 - 202 2020 - 202 2019 (5 months) 2015 (5 months)	Ranking data with continuous labels	NeurIPS, 12/201
Technology Innovation Institute (TII), Abu Dhabi, UAE AI Researcher Pompeu Fabra University, Barcelona, Spain Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem DUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202 2012 - 202 202 - 202 202 - 202 203 - 202 204 - 202 205 - 203 206 - 203 207 - 203 208 - 203 208 - 203 208 - 203 209 - 2	Max k-armed bandit: on ExtremeHunter and beyond	ECML-PKDD, 09/201
Postdoc researcher: distributional dynamic programming, safety DeepMind, London, UK PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem OUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2015 - 202 2016 - 202 2017 - 202 2017 - 202 2018 - 202 2019 (5 months) 2018 (5 months) 2019 (5 months) 2015 (5 months)		2022 - 2024
PhD internship: distributional (and deep) reinforcement learning, exploration Inria, SequeL team, Villeneuve d'Ascq, France Research internship: multi-armed bandit, shortest path problem OUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2015 - 202		2020 - 2021
Research internship: multi-armed bandit, shortest path problem OUCATION Télécom Paris, LTCI lab, Palaiseau, France PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 202		2019 (5 months)
Télécom Paris, LTCI lab, Palaiseau, France2016 - 202PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning'2015 - 202École normale supérieure de Cachan, Cachan, France2015 - 202MSc 'MVA': machine learning2012 - 202École polytechnique, Palaiseau, France2012 - 202		2015 (5 months)
PhD thesis: 'Ranking and Risk-Aware Reinforcement Learning' École normale supérieure de Cachan, Cachan, France MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2015 - 2016 2012 - 2017	DUCATION	
MSc 'MVA': machine learning École polytechnique, Palaiseau, France 2012 - 201	·	2016 - 2020
École polytechnique, Palaiseau, France 2012 - 201		2015 - 2016

TECHNICAL STRENGTHS

Programming	Python, PyTorch, TensorFlow
Languages	French (fluent), English (fluent)