Johannes **Brandstetter**

MACHINE LEARNING EXPERT · PHYSICIST · DATA ANALYST

DoB: September 25, 1989
Altenberger Straße 69, 4040 Linz, AUSTRIA

+43 (0)732 2468 4526 | brandstetter@ml.jku.at | Brandstetter | Brandstetter

Summary

I am an experienced Postdoctoral Researcher with a demonstrated history of working and publishing in (Deep) Machine Learning, Data Science and High Energy Physics. During my PhD I have been working as physicist and data scientist at the CMS experiment at the Large Hadron Collider at CERN, where I was one of the main researchers in the first direct observation of the Higgs boson decaying into pairs of fermions. Currently, I am scientific staff member at the Institut for Machine Learning at the Johannes Kepler University in Linz headed by Prof. Sepp Hochreiter. My current research topics comprise Deep Learning, Deep Reinforcement Learning, Generative Adversarial Networks, physics-inspired and physics-informed machine learning, and modeling of sequential data with variants of the Long Short-Term Memory. Furthermore, I develop new Deep Learning approaches to support the design of quantum experiments and the modeling of complex environmental systems.

Research Experience

Scientific Staff Member at Institute for Machine Learning, Johannes Kepler University, Linz   *Jul. 2018 - ongoing*
- University assistent at Institute for Machine Learning (headed by Prof. Sepp Hochreiter)
- Working areas and research interests involve Deep Learning, Deep Reinforcement Learning, Generative Adversarial Networks, physics-inspired and physics-informed machine learning, modeling of sequential data with variants of the Long Short-Term Memory

Junior Researcher at Institute of High Energy Physics, Vienna   *Nov. 2013 - May 2018*
- Member of the CMS experiment, CERN
- 4.5 years research for Doctoral and Master Thesis
- Development of machine learning tools for tau lepton reconstruction and tau lepton identification
- Development of fully data-driven background estimation method (fake-factor method)
- Development and application of physics analysis and complex statistical methods
- Development of C++ based large-scale analysis frameworks
- Collaboration with internationally highly recognized universities
- Several short- and mid-range stays at CERN
- Co-supervision of bachelor and master students

Research project at CERN   *Jul. 2013 - Aug.2013*
- Working title "*Improvement of the Communication Interface of High-Level-Trigger and Level-1 Trigger*"
- Development of a common web interface for different trigger levels at the CMS experiment

Research project at Institute of Applied Physics Vienna   *Sep. 2013 - Nov.2013*
- Working title "*Electronic Structure of Diamonds with Nitrogen Impurities*"
- Simulation of diamond lattices, statistical analysis

Education

Doctoral program in Technical Sciences Technical Physics   *Jan. 2015 - May 2018*
- TU Wien (Dr.techn./Ph.D.), with distinction
- Doctoral Thesis: "*Neutral Higgs Boson and Z Boson Decays Into Pairs of Tau Leptons with the CMS Detector*", supervisor: Jochen Schieck

Master Program Technical Physics   *Oct. 2012 - Nov. 2014*
- TU Wien (Dipl.-Ing./MSc.), with distinction
- Master Thesis: "*Understanding Muon Detection Efficiencies for Quarkonium Polarization Measurements at the Compact Muon Solenoid*", supervisor: Claudia Wulz

Bachelor Program Technical Physics *Oct. 2009 - Jul. 2012*
- TU Wien (BSc.), with distinction
- Bachelor Thesis at Stefan-Meyer Institute Vienna (Mar. - Jul. 2012): "*Design of a Polarized Hydrogen Beam*",
  supervisor: Eberhard Widmann

Bundesgymnasium Schärding *2000 - 2008*
- Matura, with distinction

Community Service *Aug. 2008 - Apr. 2009*
- Austrian Civil Service at *Rotes Kreuz* Schärding

Skills

(Deep) Machine Learning	(Deep) Reinforcement Learning, Generative Adversarial Networks, Physics-inspired and physics-informed machine learning, Modeling of sequential data with Long Short-Term Memory, Object detection and segmentation with CNNs, Contribution analysis and interpretation of networks, Support Vector Machines, Random Forest and Gradient Boosting
Data Science	Data handling, Data analysis workflow, Designing of analyses, Uncertainty treatment, Hypothesis testing, Maximum Likelihood fitting
Physics	Quantum mechanics, Field theory, Electrodynamics, Statistical physics, Relativity, Classical mechanics, Computational physics
Mathematics	Probability theory, Statistics, Linear algebra, Analysis
Programming	C++, Python, C, Fortran, Bash
Deep Learning Frameworks	PyTorch
Operation Systems	Linux, Mac OS
Tools	Git/Github, SVN, Jupyter, ROOT (CERN's generic software framework), LaTeX, Mathematica, Matlab
Language Skills	*German* (Mother tongue), *English* (Excellent level), *French* (Basic level)

Further Training

CERN School of Computing 2017 *Aug. 2017 - Sep. 2017*
- Universidad Politécnica de Madrid

MLHEP Summer School on Machine Learning in High Energy Physics *Aug. 2015*
- St. Petersburg Academic University

ATHENS Program - Information Extraction Techniques and the Semantic Web *Mar. 2014*
- Télécom ParisTech

Tutor at the Institute of Solid State Physics, TU Wien *Oct. 2012 - Jun. 2013*
- Supervising laboratory exercises

Seminar Subatomic Physics *Feb. 2012*
- University of Jyväskylä

Internships

EVGroup, St.Florian/Inn *Jul. 2012 - Sep. 2012*
- Simulating and analyzing heat dependencies in wafer bonding processes

Frauscher Sensortechnik, St.Marienkirchen/Schärding *Jul. 2011 - Aug. 2011*
- Analyzing applications of inductive railway sensor technology processes

Fill Metallbau, Hohenzell/Ried *August. 2010 - Sep. 2010*
- Various metalworking tasks

Tennis Instructor *2005 - 2012*
- Responsibility for small groups of children of various age groups

List of five most significant publications

RUDDER: Return Decomposition for Delayed Rewards

Delayed rewards are very common as they typically appear in reinforcement learning tasks with episodic or sparse rewards. Delayed rewards impose a fundamental problem onto learning and the solution is a long standing challenge in reinforcement learning. Via RUDDER, we introduce a novel reinforcement approach to overcome this delayed reward problem. The central idea of RUDDER is to make the expected future rewards equal to zero. This is achieved via reward redistribution and return decomposition via contribution analysis. RUDDER assigns credit to key events and thereby speeds up learning with delayed rewards dramatically.

- RUDDER: Return Decomposition for Delayed Rewards, Jose Arjona-Medina, Michael Gillhofer, Michael Widrich, Thomas Unterthiner, Johannes Brandstetter, Sepp Hochreiter, arXiv:1806.07857, Accepted at 2019 Conference on Advances in Neural Information Processing Systems.

Observation of the Higgs boson decay to a pair of tau leptons

This paper reports the first direct observation of Higgs boson decays into pairs of tau leptons. This paper is therefore a direct confirmation of Higgs boson Yukawa coupling to leptons, one of the big pillars in the standard model of particle physics. The measurement performed with data collected by the CMS experiment leads to an observed significance of 5.9 standard deviations, equal to the expected significance. My contributions comprise the development and evaluation of machine learning tools for tau lepton reconstruction and tau lepton identification, the treatment of the full uncertainty model and the statistical inference of the signal.

- Observation of the Higgs boson decay to a pair of tau leptons, CMS Collaboration, Phys.Lett. B779 (2018) 283-316.

Measurement of the $Z\gamma^* \rightarrow \tau\tau$ cross section in pp collisions at $\sqrt{s} = 13$ TeV and validation of τ lepton analysis techniques

This paper reports a measurement of the $Z/\gamma* \rightarrow \tau\tau$ cross section in pp collisions at 13 TeV, using data recorded by the CMS experiment. In this paper, I developed a fully data-driven background estimation method and carried out the analysis from the prepocessing of the raw detector data to the final statistical analysis. A new background estimation method was necessary since several backgrounds arise from events in which a quark or gluon jet is misidentified as tau lepton decay. These jets are notoriously difficult to model via simulation. Additionally, due to relatively small misidentification rates, large simulation samples are needed to avoid prohibitively large statistical uncertainties in the modeling. For these reasons, the fake factor method was introduced which relies almost entirely on collision data such that simulation-related uncertainties become negligible and the statistical power of the model scales directly with the collected luminosity without requiring simulation samples to match the latter. For the development of the fake factor method a new machine learning approach to the identification of tau leptons was necessary.

- Measurement of the $Z\gamma^* \rightarrow \tau\tau$ cross section in pp collisions at $\sqrt{s} = 13$ TeV and validation of τ lepton analysis techniques, CMS Collaboration, Eur. Phys. J. C. (2018) 78: 708.

Search for additional neutral Higgs bosons in the tau tau final state in proton-proton collision at $\sqrt{s} = 13$ TeV

This paper presents a search for additional neutral Higgs bosons in the di-tau final state, using data recorded by the CMS experiment. The search is performed in the context of the minimal supersymmetric extension of the standard model, and has set state of the art limits for the occurence of additional neutral Higgs bosons, which are still valid today. In this paper, I carried out the analysis from the prepocessing of the raw detector data to the final statistical analysis. I further applied the fake factor method for the first time to one of the flagship analyses of the CMS experiment. In order to deal with the extremely large dataset, I developed new statistical tools to evaluate the fake factor method in unseen kinematic regions, and came up with a new method to test artificially introduced signals in regions where no coverage of simulations was available.

- Search for additional neutral Higgs bosons in the tau tau final state in proton-proton collision at $\sqrt{s} = 13$ TeV, CMS Collaboration, JHEP 09 (2018)007.

Measurement of Higgs boson production and decay to the final state

This paper is a follow-up result on the paper **"Observation of the Higgs boson decay to a pair of tau leptons"**. The analyses was extended to a larger dataset and state of the art analysis tools were applied. Results are the most precise measurement of the inclusive cross section of the Higgs boson decay into pairs of tau tau leptons so far. For a new event categorization approach I co-developed a neural network based analysis workflow. Events were categorized by a neural network into several regions, in which a maximum likelihood fit was applied. For a correct statistical inference all uncertainties had to be propagated through the network.

- Measurement of Higgs boson production and decay to the $\tau\tau$ final state, CMS Collaboration, CMS-PAS-HIG-18-032, http://inspirehep.net/record/1726688.

Contributions to Publications and Conferences

Machine Learning Publications

- **RUDDER: Return Decomposition for Delayed Rewards**, Jose Arjona-Medina, Michael Gillhofer, Michael Widrich, Thomas Unterthiner, Johannes Brandstetter, Sepp Hochreiter, arXiv:1806.07857, Accepted at 2019 Conference on Advances in Neural Information Processing Systems.
- **Quantum Optical Experiments Modeled by Long Short-Term Memory**, Thomas Adler, Manuel Erhard, Mario Krenn, Johannes Brandstetter, Johannes Kofler, Sepp Hochreiter, Accepted at the "Machine Learning and the Physical Sciences" workshop at 2019 Conference on Advances in Neural Information Processing Systems, arXiv:1910.13804.
- **A GAN based solver of black-box inverse problems**, Michael Gillhofer, Hubert Ramsauer, Johannes Brandstetter, Bernhard Schäfl, Sepp Hochreiter, Accepted at the "Solving inverse problems with deep networks: New architectures, theoretical foundations, and applications" workshop at 2019 Conference on Advances in Neural Information Processing Systems.
- **A GAN based solver idea for derivative-free optimization problems**, Hubert Ramsauer, Johannes Brandstetter, Michael Gillhofer, Bernhard Schäfl, Sepp Hochreiter, Accepted at the "Science meets Engineering of Deep Learning" workshop at 2019 Conference on Advances in Neural Information Processing Systems.
- **Using LSTMs for climate change assessment studies on droughts and floods**, Frederik Kratzert, Daniel Klotz, Johannes Brandstetter, Pieter-Jan Hoedt, Grey Nearing, Sepp Hochreiter, Accepted at the "Tackling Climate Change with ML' workshop at 2019 Conference on Advances in Neural Information Processing Systems.

Selected Physics Publications

I was part of the CMS experiment for more than five years. Here, only publications are listed where I consider myself as a main author or to which I have contributed significantly. A complete list of all physics publication can be found in **my Google Scholar profile**.

- **PhD Thesis - Neutral Higgs Boson and Z Boson Decays into Pairs of Tau Leptons with the CMS Detector**, Johannes Brandstetter, CERN-THESIS-2018-066.
- **Search for singly produced third-generation leptoquarks decaying to a tau lepton and a b quark in proton-proton collisions at $\sqrt{s} = 13$ TeV**, CMS Collaboration, JHEP 07 (2018) 115.
- **Search for additional neutral Higgs bosons in the tau tau final state in proton-proton collision at $\sqrt{s} = 13$ TeV**, CMS Collaboration, JHEP 09 (2018)007.
- **Measurement of the $Z\gamma^* \to \tau\tau$ cross section in pp collisions at $\sqrt{s} = 13$ TeV and validation of τ lepton analysis techniques**, CMS Collaboration, Eur. Phys. J. C. (2018) 78: 708.
- **Observation of the Higgs boson decay to a pair of tau leptons**, CMS Collaboration, Phys.Lett. B779 (2018) 283-316.
- **Technical proposal for the Phase-II Upgrade of the Compact Muon Solenoid**, CMS Collaboration, CMS-TDR-15-002, CERN-LHCC-2015-010, https://cds.cern.ch/record/2020886.
- **CMS Phase II Upgrade Scope Document**, CMS Collaboration, CERN-LHCC-2015-019, https://cds.cern.ch/record/2055167.
- **Jet algorithms performance in 13 TeV data**, CMS Collaboration, CMS-PAS-JME-16-003, http://cds.cern.ch/record/2256875.
- **Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2**, CMS Collaboration, CMS-PAS-TAU-16-002, https://cds.cern.ch/record/2196972.
- **Measurement of Higgs boson production and decay to the $\tau\tau$ final state**, CMS Collaboration, CMS-PAS-HIG-18-032, http://inspirehep.net/record/1726688.
- **Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis**, ATLAS and CMS Collaboration, Phys.Lett. B779 (2018) 283-316.
- **Search for a neutral MSSM Higgs boson decaying into ττ with 12.9 fb−1 of data at $\sqrt{s}$=13 TeV**, CMS Collaboration, CMS-PAS-HIG-16-037, https://cds.cern.ch/record/2231507.
- **J/psi and psi(2S) prompt double-differential cross sections in pp collisions at 7 TeV**, CMS Collaboration, CMS-PAS-BPH-14-001, https://cds.cern.ch/record/1744914.

Talks at Conferences

- **Searches for additional Higgs bosons at CMS with Run2 data**, DIS 2018, Kobe, Japan, Apr. 2018.
- **Higgs boson results on couplings to fermions, CP parameters and perspectives for HL-LHC (ATLAS AND CMS)**, LCWS 2017, Strasbourg, France, Oct.2017.
- **Higgs boson and standard model physics in the tautau final state**, ALPS 2017, Obergurgl, Austria, Apr. 2017.
- **Tau Identification Studies for High-Luminosity LHC Upgrade**, OePG 2016, Vienna, Austria, Sep. 2016.
- **Higgs properties**, FPCP 2016, California Institute of Technology, Pasadena, USA, Jun. 2016.
- **Results and prospects for the measurement of Higgs boson decays to pairs of tau leptons with the CMS experiment**, SPG-OePG 2015, Vienna, Austria, Sep. 2015.
- **Higgs boson properties and tests of the Standard Model at CMS**, QCD 2015, Montpellier, France, Jun. 2015.
- **Measurements of Quarkonia Production Cross Section at CMS**, OePG 2015, Pöllau, Austria, Sep. 2014.

Conference Proceedings

- **Higgs boson results on couplings to fermions, CP parameters and perspectives for HL-LHC (ATLAS AND CMS)**, Johannes Brandstetter, Talk presented at the International Workshop on Future Linear Colliders (LCWS2017), Strasbourg, France, 23-27 October 2017. C17-10-23.2, arXiv:1801.07926.
- **Higgs properties**, Johannes Brandstetter, PoS(FPCP2016)025.
- **Higgs boson properties and tests of the Standard Model at CMS**, Johannes Brandstetter, Nuclear and Particle Physics Proceedings, Volumes 270-272, January-March 2016, p.222-226, https://doi.org/10.1016/j.nuclphysbps.2016.02.044 .

Teaching

Teaching Experience

- 2018: Exercises in Machine Learning: Supervised Techniques, Lecturer, Johannes Kepler University, Linz
- 2018: Lecture in Numerical and Symbolic Methods for Bioinformatics, Lecturer, Johannes Kepler University, Linz
- 2019: Exercises in Machine Learning: Unsupervised Techniques, Lecturer, Johannes Kepler University, Linz
- 2019: Exercises in Theoretical Concepts of Machine Learning, Lecturer, Johannes Kepler University, Linz
- 2019: Lecture in Machine Learning: Supervised Techniques, Lecturer, Johannes Kepler University, Linz
- 2019: Exercises in Machine Learning: Supervised Techniques, Lecturer, Johannes Kepler University, Linz
- 2019: Lecture in Hands-on AI I, Lecturer, Johannes Kepler University, Linz
- 2019: Lecture in Programming in Python I, Co-Lecturer, Johannes Kepler University, Linz

I was strongly involved in the preparation of the study program "Artificial Intelligence"

- Design and preparation of the new courses "Hands-on AI I" and "Hands-on AI II"
- Restructuring of the course "Lecture in Machine Learning: Supervised Techniques"
- Restructuring of the course "Exercises in Machine Learning: Supervised Techniques"
- Restructuring of the course "Exercises in Machine Learning: Unsupervised Techniques"
- Restructuring of the course "Exercises in Theoretical Concepts of Machine Learning"

Deep Learning lecture series (lecture+exercises) at OeAW AI Summer School   *August 2019*

- Homepage of the Summer School
- Introduction to Data Science
- Introduction to Machine Learning
- Deep Learning I
- Deep Learning II
- Deep Learning III
- Deep Learning IV
- Deep Learning V

Evaluation of teaching

- For all my university courses so far, the minimum number of evaluation forms filled by students (5) was not given, such that a statistical evaluation was not possible.
- Selected comments on the question "What did you particularly like (at the OeAW AI Summer School)":
    - *Deep Learning, Jupyter Notebooks*
    - *Like how nicely the deep learning concepts were introduced. It was very articulate*
    - *the theoretical expertise and the hands on mind set of all lectures, but especially Johannes. Seeing young and ambitious scientists like him "in action" is really inspiring*
    - *Deep learning courses were the most pratical subject*
    - *I enjoyed a lot all the tips and tricks of the trade we were given for the Deep Learning module*
    - *Pytorch, deep learning*

Supervised Master Theses

- Lukas Gruber: Analysis of Integral Probability Metrics for Generative Adversarial Networks
- Kristina Duswald: Identification and classification of microplastics in environmental samples based on fourier transform infrared spectroscopy (FTIR) using deep neural network, *ongoing*

Outreach

Scientific Blogs

- Blogpost to RUDDER: Return Decomposition for Delayed Rewards
- Interview on "How AI is changing our daily life"

Scientific Communication

- Talk at SCCH Innovationspool "Qualitätsspirale nach oben: AI verbessert sich selbst" 2018
- Scientific outreach for Institute of High Energy Physics at "European Researchers' Night: be-SCIENCEd" 2017
- Scientific outreach for Institute of High Energy Physics at "Lange Nacht der Forschung" 2014, 2016

Hackathon

- Organization of Deep Learning track at oeaw Machine Learning hackathon, Vienna, January 2019

List of publications including citations counts
(Google Scholar accessed on November 11th)

I was part of the CMS experiment for more than five years. My h-index is 88. However, this is not meaningful since at the CMS experiment every core CMS member is signing every paper. In order to make a reasonably fair comparison, I therefore listed those publications where I consider myself as a main authors or to which I have contributed significantly. This results in an h-index of 12. High-energy physics citation counts are obtained from **inspire.hep**, since inspire.hep is the standard citation tool in high energy physics. I did not list my conference proceedings or my workshop papers. A complete list of all physics publication can be found in **my Google Scholar profile**.

	Cited by	
	All	Since 2014
Citations	27695	27572
h-index	88	88
i10-index	355	354
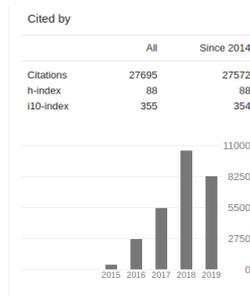
Title	Cited by	Year
Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at √=7 and 8 TeV CMS Collaboration, JHEP 1608 (2016) 045	1067	2016
Technical Proposal for the Phase-II Upgrade of the CMS Detector CMS Collaboration, CMS-TDR-15-02	354	2015
Jet algorithms performance in 13 TeV data CMS Collaboration, CMS-PAS-JME-16-003	175	2017
Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector CMS Collaboration, Physics letters B 779, 283-316	143	2018
Combined measurements of Higgs boson couplings in proton-proton collisions at 13 TeV CMS Collaboration, Eur.Phys.J. C79 (2019) no.5, 421	130	2019
Search for additional neutral MSSM Higgs bosons in the ττ final state in proton-proton collisions at s√= 13 TeV CMS Collaboration, JHEP 1809 (2018) 007	72	2018
Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2 CMS Collaboration, CMS-PAS-TAU-16-002	63	2016
Search for a neutral MSSM Higgs boson decaying into ττ with 12.9 fb−1 of data at s√=13 TeV CMS Collaboration, CMS-PAS-HIG-16-037	57	2016
Performance of reconstruction and identification of τ leptons decaying to hadrons and ντ in pp collisions at s√= 13 TeV CMS Collaboration, JINST 13 (2018) no.10, P10005	44	2018
Rudder: Return decomposition for delayed rewards JA Arjona-Medina, M Gillhofer, M Widrich, T Unterthiner, J Brandstetter, Sepp Hochreiter, arXiv preprint arXiv:1806.07857	25	2018
Search for a singly produced third-generation scalar leptoquark decaying to a τ lepton and a bottom quark in proton-proton collisions at s√=13 TeV CMS Collaboration, JHEP 07(2018) 115	21	2018
Measurement of the Z/γ∗ → ττ cross section in pp collisions at s√=13 TeV and validation of τ lepton analysis techniques CMS Collaboration, Eur. Phys. J. C 78 (2018) 708	14	2018
Measurement of the Z/γ∗ → ττ cross section in pp collisions at s√=13 TeV and validation of τ lepton analysis techniques CMS Collaboration, Eur. Phys. J. C 78 (2018) 708	5	2019