



Batch Normalization:
Accelerating Deep Network
Training by Reducing
Internal Covariate Shift

Ioffe, Szegedy
ICML 2015

Citations per year	616
Equations per page	3.00
Oldest citation*	1998
Self citations	1
Hinton references	3

*If you authored this paper
you win this round!*

**earliest year wins*



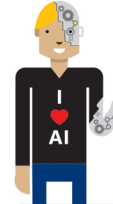
Dropout: A Simple Way to
Prevent Neural Networks
from Overfitting

*Srivastava, Hinton,
Krizhevsky, Sutskever,
Salakhutdinov*
JMLR 2014

Citations per year	640
Equations per page	0.67
Oldest citation*	1943
Self citations	11
Hinton references	7

*If you authored this paper
you win this round!*

**earliest year wins*



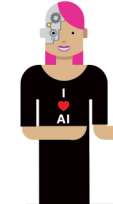
A Training Algorithm for
Optimal Margin Classifiers

Boser, Guyon, Vapnik
COLT 1992

Citations per year	554
Equations per page	2.40
Oldest citation*	1953
Self citations	6
Hinton references	0

*If you authored this paper
you win this round!*

**earliest year wins*



Support-Vector Networks

Cortes, Vapnik
Machine Learning 1995

Citations per year	1072
Equations per page	4.40
Oldest citation*	1953
Self citations	3
Hinton references	2

*If you authored this paper
you win this round!*

**earliest year wins*



Deep Residual Learning for
Image Recognition

He, Zhang, Ren, Sun
CVPR 2016

Citations per year	521
Equations per page	0.16
Oldest citation*	1989
Self citations	5
Hinton references	3

*If you authored this paper
you win this round!*

**earliest year wins*



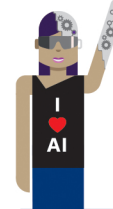
Fully Convolutional
Networks for Semantic
Segmentation

Long, Shelhamer, Darrell
CVPR 2015

Citations per year	765
Equations per page	0.33
Oldest citation*	1987
Self citations	5
Hinton references	1

*If you authored this paper
you win this round!*

**earliest year wins*



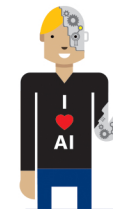
ImageNet Classification with
Deep Convolutional Neural
Networks

Krizhevsky, Sutskever, Hinton
NIPS 2012

Citations per year	1752
Equations per page	0.44
Oldest citation*	1990
Self citations	5
Hinton references	3

*If you authored this paper
you win this round!*

**earliest year wins*



Visualizing and
Understanding
Convolutional Networks

Zeiler, Fergus
ECCV 2014

Citations per year	582
Equations per page	0.00
Oldest citation*	1989
Self citations	2
Hinton references	3

*If you authored this paper
you win this round!*

**earliest year wins*





Microsoft



Microsoft



Microsoft



Microsoft



Microsoft



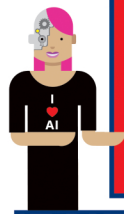
Microsoft



Microsoft



Microsoft



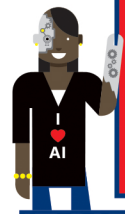
Context-Dependent
Pre-Trained Deep Neural
Networks for Large-
Vocabulary Speech
Recognition

Dahl, Yu, Deng, Acero
IEEE Trans. Audio 2012

Citations per year	291
Equations per page	2.25
Oldest citation*	1986
Self citations	21
Hinton references	13

*If you authored this paper
you win this round!*

*earliest year wins



Practical Bayesian
Optimization of Machine
Learning Algorithms

Snoek, Larochelle, Adams
NIPS 2012

Citations per year	113
Equations per page	1.00
Oldest citation*	1964
Self citations	1
Hinton references	0

*If you authored this paper
you win this round!*

*earliest year wins



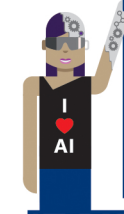
Keeping Neural Networks
Simple by Minimizing the
Description Length of the
Weights

Hinton, van Camp
COLT 1993

Citations per year	17
Equations per page	2.08
Oldest citation*	1986
Self citations	2
Hinton references	2

*If you authored this paper
you win this round!*

*earliest year wins



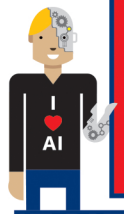
A Practical Bayesian
Framework for
Backpropagation Networks

MacKay
Neural Computation 1992

Citations per year	73
Equations per page	0.88
Oldest citation*	1967
Self citations	2
Hinton references	2

*If you authored this paper
you win this round!*

*earliest year wins



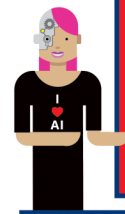
Sequence to Sequence
Learning with Neural
Networks

Sutskever, Vinyals, Le
NIPS 2014

Citations per year	422
Equations per page	0.56
Oldest citation*	1986
Self citations	2
Hinton references	3

*If you authored this paper
you win this round!*

*earliest year wins



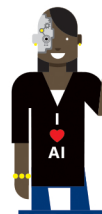
Human-level Control
through Deep
Reinforcement Learning

*Mnih, Kavukcuoglu, Silver,
Rusu, Veness et al.*
Nature 2015

Citations per year	499
Equations per page	0.54
Oldest citation*	1911
Self citations	6
Hinton references	4

*If you authored this paper
you win this round!*

*earliest year wins



The Helmholtz Machine

Dayan, Hinton, Neal, Zemel
Neural Computation 1995

Citations per year	34
Equations per page	1.58
Oldest citation*	1956
Self citations	9
Hinton references	6

*If you authored this paper
you win this round!*

*earliest year wins



A Bayesian Hierarchical
Model for Learning Natural
Scene Categories

Fei-Fei, Perona
CVPR 2005

Citations per year	283
Equations per page	1.50
Oldest citation*	1980
Self citations	1
Hinton references	0

*If you authored this paper
you win this round!*

*earliest year wins





Microsoft



Microsoft



Microsoft



Microsoft



Microsoft



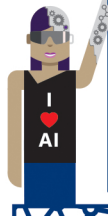
Microsoft



Microsoft



Microsoft



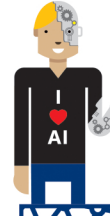
Training with Noise is
Equivalent to Tikhonov
Regularization

Bishop
Neural Computation 1995

Citations per year	24
Equations per page	4.00
Oldest citation*	1977
Self citations	3
Hinton references	0

If you authored this paper
you win this round!

*earliest year wins



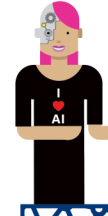
Latent Dirichlet Allocation

Blei, Ng, Jordan
JMLR 2003

Citations per year	1240
Equations per page	1.90
Oldest citation*	1970
Self citations	3
Hinton references	0

If you authored this paper
you win this round!

*earliest year wins



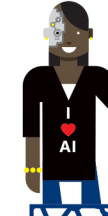
Loopy Belief Propagation
for Approximate Inference:
An Empirical Study

Murphy, Weiss, Jordan
UAI 1999

Citations per year	84
Equations per page	0.78
Oldest citation*	1988
Self citations	3
Hinton references	0

If you authored this paper
you win this round!

*earliest year wins



A Global Geometric
Framework for Nonlinear
Dimensionality Reduction

*Tenenbaum, de Silva,
Langford*
Science 2000

Citations per year	596
Equations per page	1.75
Oldest citation*	1970
Self citations	1
Hinton references	1

If you authored this paper
you win this round!

*earliest year wins



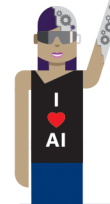
Long Short-Term Memory

Hochreiter, Schmidhuber
Neural Computation 1997

Citations per year	137
Equations per page	2.02
Oldest citation*	1987
Self citations	10
Hinton references	1

If you authored this paper
you win this round!

*earliest year wins



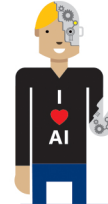
Auto-Encoding Variational
Bayes

Kingma, Welling
ICLR 2014

Citations per year	191
Equations per page	3.36
Oldest citation*	1986
Self citations	0
Hinton references	1

If you authored this paper
you win this round!

*earliest year wins



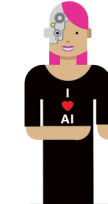
A View of The EM Algorithm
that Justifies Incremental,
Sparse, and Other Variants

Neal, Hinton
Learning in Graphical
Models 1998

Citations per year	128
Equations per page	1.36
Oldest citation*	1977
Self citations	0
Hinton references	0

If you authored this paper
you win this round!

*earliest year wins



Max-Margin Markov
Networks

Taskar, Guestrin, Koller
NIPS 2004

Citations per year	108
Equations per page	2.38
Oldest citation*	1988
Self citations	1
Hinton references	0

If you authored this paper
you win this round!

*earliest year wins





Microsoft



Microsoft



Microsoft



Microsoft



Microsoft



Microsoft



Microsoft



Microsoft