

# Where is Machine Learning Going?

**Francis Bach**

*INRIA - Ecole Normale Supérieure, Paris, France*



NeurIPS Workshop on Experiment Pre-Registration  
*November 2020*

# Where is Machine Learning Going?

**A personal and biased view**

# Where is Machine Learning Going?

## A personal and biased view

- **NeurIPS: from 2000 to 2020**
- **Shift in values**
- **A scientific approach to machine learning**
- **BONUS: Should my daughter do a PhD in machine learning?**

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
contested US election results	contested US election results

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
contested US election results 300 participants	contested US election results 10 000 participants

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p data-bbox="163 326 966 381">contested US election results</p> <p data-bbox="289 428 842 578">300 participants (very) little diversity</p>	<p data-bbox="1066 326 1864 381">contested US election results</p> <p data-bbox="1199 428 1730 578">10 000 participants little diversity</p>

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p data-bbox="163 326 968 380">contested US election results</p> <p data-bbox="289 428 842 574">300 participants (very) little diversity</p> <p data-bbox="472 618 659 672">kernels</p>	<p data-bbox="1066 326 1864 380">contested US election results</p> <p data-bbox="1199 428 1730 574">10 000 participants little diversity</p> <p data-bbox="1251 618 1677 672">neural networks</p>



# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p data-bbox="170 329 961 378">contested US election results</p> <p data-bbox="289 431 842 578">300 participants (very) little diversity</p> <p data-bbox="474 618 657 667">kernels</p> <p data-bbox="306 724 825 773">short 7-page paper</p>	<p data-bbox="1066 329 1858 378">contested US election results</p> <p data-bbox="1199 431 1730 578">10 000 participants little diversity</p> <p data-bbox="1251 618 1677 667">neural networks</p> <p data-bbox="1024 724 1902 773">20-page supplementary material</p>

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p>contested US election results</p> <p>300 participants (very) little diversity</p> <p>kernels</p> <p>short 7-page paper toy simulations</p>	<p>contested US election results</p> <p>10 000 participants little diversity</p> <p>neural networks</p> <p>20-page supplementary material large-scale experiments</p>

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p data-bbox="170 329 961 378">contested US election results</p> <p data-bbox="289 431 842 574">300 participants (very) little diversity</p> <p data-bbox="474 618 659 667">kernels</p> <p data-bbox="306 724 825 948">short 7-page paper toy simulations small computing power</p>	<p data-bbox="1066 329 1858 378">contested US election results</p> <p data-bbox="1199 431 1730 574">10 000 participants little diversity</p> <p data-bbox="1251 618 1677 667">neural networks</p> <p data-bbox="1024 724 1900 948">20-page supplementary material large-scale experiments lots of computing power</p>

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p>contested US election results</p> <p>300 participants (very) little diversity</p> <p>kernels</p> <p>short 7-page paper</p> <p>toy simulations</p> <p>small computing power</p> <p>Matlab</p>	<p>contested US election results</p> <p>10 000 participants little diversity</p> <p>neural networks</p> <p>20-page supplementary material</p> <p>large-scale experiments</p> <p>lots of computing power</p> <p>Python</p>

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p>contested US election results</p> <p>300 participants (very) little diversity</p> <p>kernels</p> <p>short 7-page paper</p> <p>toy simulations</p> <p>small computing power</p> <p>Matlab</p> <p>few preprints</p>	<p>contested US election results</p> <p>10 000 participants little diversity</p> <p>neural networks</p> <p>20-page supplementary material</p> <p>large-scale experiments</p> <p>lots of computing power</p> <p>Python</p> <p>ArXiv</p>

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p>contested US election results</p> <p>300 participants (very) little diversity</p> <p>kernels</p> <p>short 7-page paper</p> <p>toy simulations</p> <p>small computing power</p> <p>Matlab</p> <p>few preprints</p> <p>tiny industry presence</p>	<p>contested US election results</p> <p>10 000 participants little diversity</p> <p>neural networks</p> <p>20-page supplementary material</p> <p>large-scale experiments</p> <p>lots of computing power</p> <p>Python</p> <p>ArXiv</p> <p>massive industry presence</p>

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p>contested US election results</p> <p>300 participants (very) little diversity</p> <p>kernels</p> <p>short 7-page paper</p> <p>toy simulations</p> <p>small computing power</p> <p>Matlab</p> <p>few preprints</p> <p>tiny industry presence</p> <p>no money</p>	<p>contested US election results</p> <p>10 000 participants little diversity</p> <p>neural networks</p> <p>20-page supplementary material</p> <p>large-scale experiments</p> <p>lots of computing power</p> <p>Python</p> <p>ArXiv</p> <p>massive industry presence</p> <p>lots of money</p>

# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p>contested US election results</p> <p>300 participants (very) little diversity</p> <p>kernels</p> <p>short 7-page paper</p> <p>toy simulations</p> <p>small computing power</p> <p>Matlab</p> <p>few preprints</p> <p>tiny industry presence</p> <p>no money</p> <p>nobody cared</p>	<p>contested US election results</p> <p>10 000 participants little diversity</p> <p>neural networks</p> <p>20-page supplementary material</p> <p>large-scale experiments</p> <p>lots of computing power</p> <p>Python</p> <p>ArXiv</p> <p>massive industry presence</p> <p>lots of money</p> <p>lot of (external) attention</p>



# NeurIPS: from 2000 to 2020

NIPS 2000	NeurIPS 2020
<p>contested US election results</p> <p>300 participants (very) little diversity</p> <p>kernels</p> <p>short 7-page paper</p> <p>toy simulations</p> <p>small computing power</p> <p>Matlab</p> <p>few preprints</p> <p>tiny industry presence</p> <p>no money</p> <p>nobody cared</p> <p>no social media</p>	<p>contested US election results</p> <p>10 000 participants</p> <p>little diversity</p> <p>neural networks</p> <p>20-page supplementary material</p> <p>large-scale experiments</p> <p>lots of computing power</p> <p>Python</p> <p>ArXiv</p> <p>massive industry presence</p> <p>lots of money</p> <p>lot of (external) attention</p> <p>Twitter</p>

**Shift in values**

**NB: It is not all bad!**

## Shift in values

**NB: It is not all bad!**

- More professionalism

## Shift in values

**NB: It is not all bad!**

- More professionalism
- Strong industry presence

## Shift in values

**NB: It is not all bad!**

- More professionalism
- Strong industry presence
- Hype, social networks and marketing

# Shift in values

**NB: It is not all bad!**

- **More professionalism**
- **Strong industry presence**
- **Hype, social networks and marketing**
- **Some wisdom from physics:**

*Physical Review adheres to the following policy with respect to use of terms such as “new” or “novel:” All material accepted for publication in the Physical Review is expected to contain new results in physics. Phrases such as “new,” “for the first time,” etc., therefore should normally be unnecessary; they are not in keeping with the journal’s scientific style. Furthermore, such phrases could be construed as claims of priority, which the editors cannot assess and hence must rule out.*

## Shift in values

**NB: It is not all bad!**

- More professionalism
- Strong industry presence
- Hype, social networks and marketing
- Increased competition

# Shift in values

**NB: It is not all bad!**

- More professionalism
- Strong industry presence
- Hype, social networks and marketing
- Increased competition
- From science to engineering



# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal

# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal
- Level 0: rigorous reporting of results (error bars, **true testing set**)

# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal
- Level 0: rigorous reporting of results (error bars, **true testing set**)
- Level 1: ablation studies (what is important and generalizable?)

# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal
- Level 0: rigorous reporting of results (error bars, **true testing set**)
- Level 1: ablation studies (what is important and generalizable?)
- Level 2: pre-registration of experiments

# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal
- Level 0: rigorous reporting of results (error bars, **true testing set**)
- Level 1: ablation studies (what is important and generalizable?)
- Level 2: pre-registration of experiments

- **Theoretical research**

- Learning theory can lack rigor as well

# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal
- Level 0: rigorous reporting of results (error bars, **true testing set**)
- Level 1: ablation studies (what is important and generalizable?)
- Level 2: pre-registration of experiments

- **Theoretical research**

- Learning theory can lack rigor as well
- Level 0: correct proofs

# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal
- Level 0: rigorous reporting of results (error bars, **true testing set**)
- Level 1: ablation studies (what is important and generalizable?)
- Level 2: pre-registration of experiments

- **Theoretical research**

- Learning theory can lack rigor as well
- Level 0: correct proofs
- Level 1: consistent / relevant set of assumptions

# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal
- Level 0: rigorous reporting of results (error bars, **true testing set**)
- Level 1: ablation studies (what is important and generalizable?)
- Level 2: pre-registration of experiments

- **Theoretical research**

- Learning theory can lack rigor as well
- Level 0: correct proofs
- Level 1: consistent / relevant set of assumptions
- Level 2: matching theory and practice



# A scientific approach to machine learning

- **Empirical research**

- Minimizing “test error” on a benchmark is not the only goal
- Level 0: rigorous reporting of results (error bars, **true testing set**)
- Level 1: ablation studies (what is important and generalizable?)
- Level 2: pre-registration of experiments

- **Theoretical research**

- Learning theory can lack rigor as well
- Level 0: correct proofs
- Level 1: consistent / relevant set of assumptions
- Level 2: matching theory and practice

- **Many interesting research questions beyond benchmarks**

**Should my daughter do a PhD in machine learning?**

# Should my daughter do a PhD in machine learning?

- **Disclaimer: whatever I say, she does the opposite...**

# Should my daughter do a PhD in machine learning?

- **Disclaimer: whatever I say, she does the opposite...**
- **Stress and competition**
  - ArXiv and deadlines

# Should my daughter do a PhD in machine learning?

- **Disclaimer: whatever I say, she does the opposite...**
- **Stress and competition**
  - ArXiv and deadlines
- **Why do we do scientific research for?**
  - To save the world
  - As a game
  - To make money?
  - etc..

# Should my daughter do a PhD in machine learning?

- **Disclaimer: whatever I say, she does the opposite...**
- **Stress and competition**
  - ArXiv and deadlines
- **Why do we do scientific research for?**
  - To save the world
  - As a game
  - To make money?
  - etc..
- **Where is machine learning going?**