Course Organization:

Neuroimaging and Machine Learning for Biomedicine

Intro
Lecture

Maxim Sharaev
RAIC, ISIP group

About myself

Maxim Sharaev, Ph.D.

- *2006—2012* Lomonosov MSU, BioChemical Physics, Dept. of Physics
- *2009—now* Istitute of Higher Nervous Activity and Neurophysiology, RAS
- *2012—2016* Lomonosov MSU, Ph.D. BioChemical Physics
- *2014—2018* NRC “Kurchatov Institute” Neuroimaging methods / MRI
- *2018—now* Skoltech, Machine Learning in Neuroscience and neuroimaging
- **Core:** biophysics, statistics and data analysis, mathematical modelling

Skoltech
Skolkovo Institute of Science and Technology

Agenda

- Why study neuroimaging at a Data Science program?
- Course outline
- Course project
- Course assessment

Why study neuroimaging at a Data Science program?

Neuroscience is cool! ☺

From sciencephoto.com

- Different research questions:
  - ✔ how the brain works
  - ✔ why we behave in one way or another
  - ✔ can we decode our thoughts…
- Different applied questions:
  - ✔ can we build artificial neural systems?
  - ✔ how can we communicate with brain?
  - ✔ can we effectively diagnose severe brain diseases and help the doctor to treat them?

Why study neuroimaging at a Data Science program?

- Unique opportunities to **collect measurements**
- We measure extremely complex processes and thus get **complex data**
- All this requires **skills in data analysis + knowing the data**
- In recent years we add **ML to neuroscientific problems** – so that we can built predictive models, CV systems, …

Why study neuroimaging at a Data Science program?

An increasing number of **startups** dedicated to **ML in biomedicine** with **focus on neuroscience**, see for example: *sbermed.ai* *neuralink.com* *ira-labs*

- Are missing ML engineers with **background in neuro**
- Have to invent their unique solutions – **lack of standardization**

AI systems in medicine: current state

FDA-Registered Facilities

Program	Domestic	Foreign	Total
Animal Drugs	1,608	1,156	2,764
Animal Food	17,259	6,895	24,154
Biologics	4,832	439	5,271
Human Drugs	3,647	4,159	7,806
Human Food	80,525	108,998	189,523
Medical Devices	13,790	12,891	26,681
Tobacco	3,371	0	3,371
Total	**125,032**	**134,538**	**259,570**

***IMDRF/SaMD WG/N41 (PF): Clinical Evaluation: Software as a Medical Device (SAMD)**
(U.S. Food and Drug Administration, FDA)

The number of **life science** papers describing **AI/ML**:

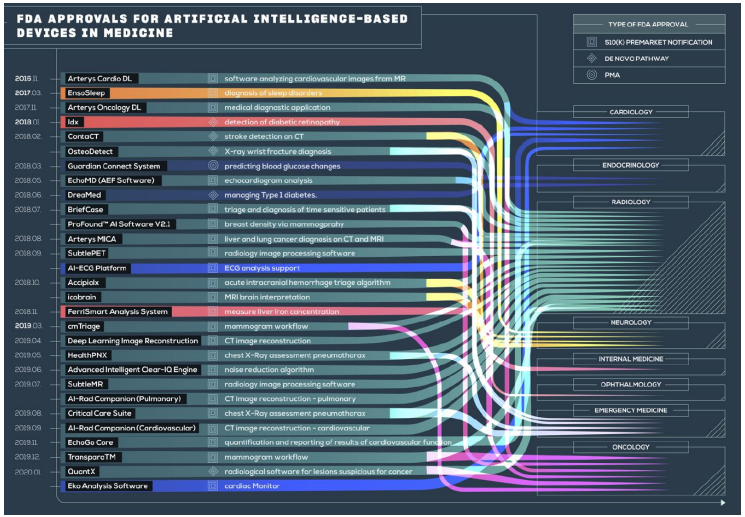
596 in 2010

12,422 in 2019

Only **64 AI/ML based**, FDA approved medical devices and algorithms

Only **29 (45%) with real AI/ML**

Less than ~0,2% of all medical devices



* from (Benjamens et al., 2020)

Skoltech
Skolkovo Institute of Science and Technology

Course outline

Goals of this course:

- Provide an introduction into the **ideas** behind neuroimaging data analysis
  - ✔ equipment, tasks and methods, data peculiarities, ML in neuroscience, …
- Learn the **tools** for neuroimaging data processing and analysis:
  - ✔ Nilearn, Freesurfer, MNE, …
- Gain the **skills** needed to solve applied clinical and scientific problems
  - ✔ Working with data processing pipelines and neuroimaging databases, creating algorithmic chains (i.e. Docker images)
- Immerse into **real clinical processes** with leading professionals from the field

What you will not learn:

ML methods *per se* (except for a narrow subset)

Course outline

Course structure for 2021:

2 parts

- Part 1: Neuroimaging experiments, methods, equipment and data analysis (6 lectures + 7 seminars)

- Part 2: Work in teams on the Final Project, panel sessions with clinicians/neuroscientists (1 seminar + QA sessions + 2-3 panel sessions)

Course outline

Classes at 12.30-15.30	Week 1 (5.09, 9.09)	Week 2 (12.09, 16.09)	Week 3 (19.09, 23.09)	Week 4 (26.09, 30.09)	Week 5 (03.10, 07.10)	Week 6 (10.10, 14.10)	Week 7 (17.10, 21.10)	Week 8 (24.10, 28.10)
Day 1, Monday. Lectures pre-recorded or online classes	Lecture 0 sync + Lecture 1 (Intro)	Lecture 3 (MRI) + **Bonus Lecture 4* (MEG)**	Lecture 5 (fMRI 1) + **Quiz fMRI-1**	Seminar 5 (fMRI ML)	Seminar 7 (Interpretation)	Work on Final Project	Panel 2 (EEG at 16.00, online)	Panel Results Discussion
Day 1, Monday. Seminar	Seminar 0 (bash, logging, data access)	Seminar 1 (EEG), part 2 Class Room: R2-B5-202 7	Seminar 3 (MRI ML)	Q/A + HW at class	Work on Final Project (prepare a plan)	Work on Final Project	Work on Final Project	Work on Final Project
		Announce: FSDocker install for Seminar 2	Announce: DockerFile install for Seminar 4					
				HW1 deadline 28.(	HW2 deadline 5.10	HW3 deadline 12.10		
Day 2, Friday. Lectures pre-recorded or online classes	Lecture 2 (EEG) + **Quiz Intro, Quizz EEG**	Lecture 3 (MRI intro) + **Quiz MRI**	Leture 6 (fMRI 2) + **Quiz fMRI-2**	Project proposals	Panel sessions presentation	Panel 1 (MRI at 17-30, offline)	Project consultations	Final projects defence
Day 2, Friday. Seminar	Seminar 1 (EEG), part 1	Seminar 2 (MRI)	Seminar 4 (fMRI)	Seminar 6 (fMRI DL)	Work on Final Project (check the plan)	Work on Final Project	Homeworks Q/A, appeals	
	Announce: HW1 (kaggle)	Announce: HW2			Announce: HW3	Project proposals review		

TBA - to be announced

Activity, attendance accounted

Assignments

Announcements

Work on Final Project, attendance accounted

Skoltech
Skolkovo Institute of Science and Technology

Course outline

Structure of a Part 1 typical learning day:

- Pre-recorded lecture [60—90 min]
- Offline quiz [should take 15 min, is closed **5 min before** the seminar]
- Offline live seminar [~80 min], **attendance is necessarily (70%)**
- HW announcement [totally 3 HWs] → submit for assessment

Part 2

- [Week 5] Project proposals [120–180 min] → submit for assessment [Week 6]
- 2-3 panel sessions: **watch** the specialist work **and propose** a solution
- Work on project → final project + peer review [Week 8]

NB! Can be flexible, follow updates in Canvas/TG Group

Course project

One of the main educational format used in this course

- **Goal:** create your own end-to-end AI system for biomedical use (see the available topics)
  - ✔ IN: lots of Neuroimaging/ML papers, datasets
  - ✔ OUT: an intelligent system for prediction / treatment planning / …
- Project is performed individually or in teams of 2~3 people with distinct roles (e.g., requirements/data analyst, data engineer, researcher, developer, …)
- Each team implements a particular data processing pipeline (e.g., data collection, cleaning, ML model validation, …)

Course assessment

Total assessment:

20% - quizzes*

45% - HW

30+5% - Final Project*****

Bonus scores:

15% Panel with clinicians

*!Quizzes are counted only if the student is present at the class!
**If there is a paper draft
***(30% - Proposal 70% - final defense)
Participation in the discussion of the progress with TA during “Work on Final Project” activity is mandatory!

Questions?

Let’s begin!