2018 Spring PRISME Forum Technical Meeting Theme:

Artificial Intelligence in Healthcare Part II: <u>The Practical Application</u>

A European Perspective



PRISME Forum – Technical Meeting

Artificial Intelligence in Healthcare Part II: The Practical Application – A European Perspective

May, 2018, Paris

Setting the Stage for the Spring PRISME Forum TM: Converging Developments Fueling the Application of AI in Pharma R&D

Juergen Hammer, PhD, MBA Global Head Data Science Roche Pharma Research and Early Development Informatics (pREDi) Roche New York





> A personal anecdote: AI example in Immunology

- > Converging developments: What is driving AI in Drug Discovery?
- Practical applications: AI/DL impacting Pharma R&D w/ reference to agenda

Conclusions



A Personal Anecdote: MHC II Epitope Repertoire Prediction & AI





The Goal: Prediction of genome-wide MHC II binding peptide repertoires presented to CD4 T cells



Predicting MHC II Peptide Repertoires:

"AI" Twenty Years ago

Al Expert System



Tiziana Sturniolo¹, Elisa Bono¹, Jiayi Ding¹, Laura Raddrizzani², Oezlem Tuerech², Juger Sahin³, Michael Braxenthaler², Fabio Gallazzi¹, Maria Pia Prott⁴, Francesco Sinigaglia¹, and Juergen Hammer^{1,2}* Recht Milan Richer, 2013 Milan, July ²Department Generalic and Information Sciences, Hoffman La Robe, NUOTIO, ¹Department of Interna Medicine, University of Sanland, 66(2) Howhang, Germany, ¹Laboratory of Tumor Immovingo, Scientific Institute II. San Righed, 2013 Milan, Italy, ¹Corresponding autor (et al.), ¹Corport, June 10, 100 (2014), ¹Corport, ¹Corpor

[Nature Biotechnology 1999]

AI Machine Learning (ANN)

Prediction of MHC class II-binding peptides using an evolutionary algorithm and artificial neural network

Vladimir Brusic, George Rudy, Margo Honeyman, Jürgen Hammer¹ and Leonard Harrison

The Walter and Eliza Hall Institute of Medical Research, PO Royal Melbourne Hospital, Parkville, Victoria 3050, Australia and ¹Roche Milano Richerche, Milan, Italy

Received on May 19, 1997, revised on August 23, 1997, accepted on August 25, 1997

[Bioinformatics 1998]





1999 News "humanizes computers"...

- Need a vaccine? Just ask a computer [Chemical & Engineering News]
- Computers Aid Vaccine Design [Science]
- Calculating Immunity [Scientific American]



AI Application in Biopharma since 20+ years Machine Learning Examples...

0

NON-TOX

00



Early 2000s... Toxicogenomic Support Vector Machine (SVM)



From 2000... Disease Molecular Taxonomy Unsupervised Clustering

1990s... Lead Optimization (QSAR, Tox,...) LR, Random Forest, SVM, shallow ANNs



Late 1990s... Epitope Predictions ANN, SVM, ES



AI Definition...? **Three Main Categories**





"A computer would deserve to be called intelligent if it could deceive a human into believing that it was



Pharmaceutical RSD Information Systems Management Executive Forum

PRISME TM Fall 2017 adopted IBM definition in white paper:

By AI we mean anything that makes machines act more intelligently

Artificial Intelligence (AI)

Expert System [Rule-based; Hard-coded]

Machine Learns [Handpicked Features]

Machine Learns How to Learn

[Automatic Learning of Features]

Classical ML; handdesigned features, e.g.

- Linear regression
- Logistic regression
- Random Forest
- Support Vector Machines
- PCA
- Clustering
- Shallow ANNs

٠ ...

Interlude Quick Hands-up exercise...





➤ <u>TM Audience question</u>:

1) Will "AI in R&D" evolve continuously like during the last 20 years?

...or...

2) Are we witnessing a true "AI in R&D" step change, as media seems to indicate?





- > A personal anecdote: AI example in Immunology
- Converging developments: What is driving AI in Drug Discovery?
- Practical applications: AI/DL impacting Pharma R&D w/ reference to agenda
- Conclusions



Awakening of AI in Biopharma as result of converging developments

- 1) Advances in **compute power**
- 2) Deep Learning **software packages** enable large community of computational scientist
- 3) Deep Neural Networks evolved & penetrated consumer products
- 4) Investments in the **Data Science** workforce
- 5) Availability of larger biomedical data sets enables training of DNNs
- 6) Industry 4.0 fear leading to investments in Biopharma AI capabilities





Awakening of AI in Biopharma as result of converging developments

- 1) Advances in compute power
- 2) Deep Learning **software packages** enable large community of computational scientist
- 3) Deep Neural Networks evolved & penetrated consumer products
- 4) Investments in the **Data Science** workforce
- 5) Availability of larger biomedical data sets enables training of DNNs
- 6) Industry 4.0 fear leading to investments in Biopharma AI capabilities





Awakening of AI in Biopharma as result of converging developments



Advances in compute power

- Deep Learning software packages enable large community of computational scientist
- 3) Deep Neural Networks evolved & penetrated consumer products
- 4) Investments in the **Data Science** workforce
- 5) Availability of larger biomedical data sets enables training of DNNs
- 6) Industry 4.0 fear leading to investments in Biopharma AI capabilities







Expert System [Rule-based; Hard-coded]

Machine Learns [Handpicked Features]

Machine Learns How to Learn

[Automatic Learning of Features]

How does

Biopharma

benefit?

Supervised Deep Learning requires:

Large Data Sets for automatic feature learning

Ground Truth!









Supporting Image-based Biomarker





Data Categories	R&D Impact	
Human Activity (Sensor) Data	Diseases w/ movement disorders; wearables/robotic biomarker	
		V

Wearable Sensors

AI & Sensor/Robotic Biomarkers

- Greater consistency...
- Higher study efficiency...
- Better Clinical endpoints?!



Robotic Sensors





Dimitris (Covance)

< Ro	che >
/	

Data Categories	R&D Impact	
Human Activity (Sensor) Data	Diseases w/ movement disorders; wearables/robotic biomarker	\checkmark
Image Data	Biomarkers; process improvements; Dx	\checkmark

AI & Imaging

- Clinical MRI, pathology, process improvement, biomarker discovery, Dx
- Improvement of efficiency and reproducibility
- Transfer Learning & Data Augmentation proven strategy to cope with small training sets



AI/DNN applications in R&D – Emerging examples...

Data Categories	R&D Impact
Human Activity (Sensor) Data	Diseases w/ movement disorders; wearables/robotic biomarker
Image Data	Biomarkers; process improvements; Dx
RWD, EHR	Clinical predictions (e.g. survival); trial design (patient recruitment, cohort matching, virtual arms)





Koroush

(Inato)

AI & RWD

- Clinical prediction, trial design & recruitment
- Scarcity of 'labels' biggest obstacle of EHR-based analysis employing ML/DL
- Deep Patient: Unsupervised DL denoising Auto-Encoder for better patient representation?





< Ro	che >

Data Categories	R&D Impact	
Human Activity (Sensor) Data	Diseases w/ movement disorders; wearables/robotic biomarker	V
Image Data	Biomarkers; process improvements; Dx	V
RWD, EHR	Clinical predictions (e.g. survival); trial design (patient recruitment, cohort matching, virtual arms)	V
Genetics, Genomics Data, Single-cell NGS	Better disease understanding (molecular & system level), back- translation (targets/pathways)	V

Fabian (Helmholtz)

AI & (SC)NGS

- Multi-omics integration w/ Deep Learning
- Auto-Encoders: Unsupervised feature construction & interpretation
- Prediction of clinical features, relevant genes, pathways
- DeepSEA: Predicting effects of non-coding variants





Data Categories	R&D Impact	
Human Activity (Sensor) Data	Diseases w/ movement disorders; wearables/robotic biomarker	Ŀ
Image Data	Biomarkers; process improvements; Dx	
RWD, EHR	Clinical predictions (e.g. survival); trial design (patient recruitment, cohort matching, virtual arms)	
Genetics, Genomics Data, Single-cell NGS	Better disease understanding (molecular & system level), back- translation (targets/pathways)	Ŀ
Compounds	Lead Optimization, Computer Aided Drug Design	[

AI & LI/LO

- > DNNs for QSAR, Tox, ADME prediction
- Generative DL models for Chemical Design
- AI to design, test and optimize compounds in silico as part of LO?

Chemical Design





Adrian Mathias (Dassault Systemes)





Yann (Iktos)

Quentin

(Iktos)

AI applications in R&D processes – Emerging examples...



- WF digitalization likely catalyst for many future AI applications
- Pharmacovigilance more scalable; automating CSR writing; RPA implementation



Awakening of AI in Biopharma as result of converging developments

Advances in compute power

- Deep Learning software packages enable large community of computational scientist



- **Deep Neural Networks** evolved & penetrated consumer products
- 4) Investments in the **Data Science** workforce
- 5) Availability of larger biomedical data sets enables training of DNNs
- 6) Industry 4.0 fear leading to investments in Biopharma AI capabilities









The Roche Advanced Analytics Data (RAAD) Experiment...

- Group-wide Challenge reveals strong Al capabilities beyond DS Functions
- Multiple AI approaches applied; decision trees (XGBoost) effective methodological choice
- A demonstration that large existing Alcapable workforce can be mobilized via Network and Challenge





[Science, May 2018]

- …machine learning algorithms … have become a form of alchemy…
- …AI researchers training their algorithms are simply stumbling in the dark…
- Al's Reproducibility problem → inconsistent experimental and publication praxis
- ii. Al's "Black Box" problem → difficulty of explaining how an AI model came to conclusion
- iii. Al's "Alchemy" problem → Data Scientists' lack of understanding of basic tools required to train algorithms; trial & error

Does Biopharma sufficiently invest in Data Science capabilities required to cope with AI/DL advances?



Awakening of AI in Biopharma as result of converging developments



Advances in compute power

- Deep Learning **software packages** enable large community of computational scientist



Deep Neural Networks evolved & penetrated consumer products



- Investments in the Data Science workforce
- 5) Availability of larger biomedical data sets enables training of DNNs
- 6) Industry 4.0 fear leading to investments in Biopharma AI capabilities



Roche

DNNs are still data-hungry in Biopharma...

- Leverage Transfer Learning (e.g. image; human activity)
- Premium payments for curated data access (e.g. EHR)
- Insufficient data access & integration of clinical/highdim/omics data/RWD (system monoliths)
- More precompetitive alliances or consortia to pre-train DNNs w/ combined data?





Awakening of AI in Biopharma as result of converging developments



Advances in compute power



Deep Learning **software packages** enable large community of computational scientist



Deep Neural Networks evolved & penetrated consumer products



Investments in the Data Science workforce



Availability of larger biomedical data sets enables training of DNNs



Industry 4.0 fear leading to investments in Biopharma AI capabilities





- > A personal anecdote: AI example in Immunology
- Converging developments: What is driving AI in Drug Discovery?
- > Practical applications: AI/DL impacting Pharma R&D w/ reference to agenda

Conclusions





Rich AI (ML) history in Biopharma R&D over past 2 decades

- DL + Compute (GUPs) + Ease-of-use (DL Software) → STEP CHANGE of AI in consumer market
- As for Biopharma, no step change yet, but many promising AI/DL applications; We will see a wide range TODAY!
- Biopharma struggle w/ sufficient Ground Truth Data for DNNs

One more time: Network, Learn and Have Fun today!