AI for Safety Assessment

Joshua Xu, Ph.D

Disclaimer: I give this presentation in my personal capacity. The views expressed are of the presenter and do not necessarily represent the views of the FDA or the United States.

Outline

- The evolution of AI through the lens of image analysis
  - Feature engineering in the early era
  - Deep learning
  - Foundational models

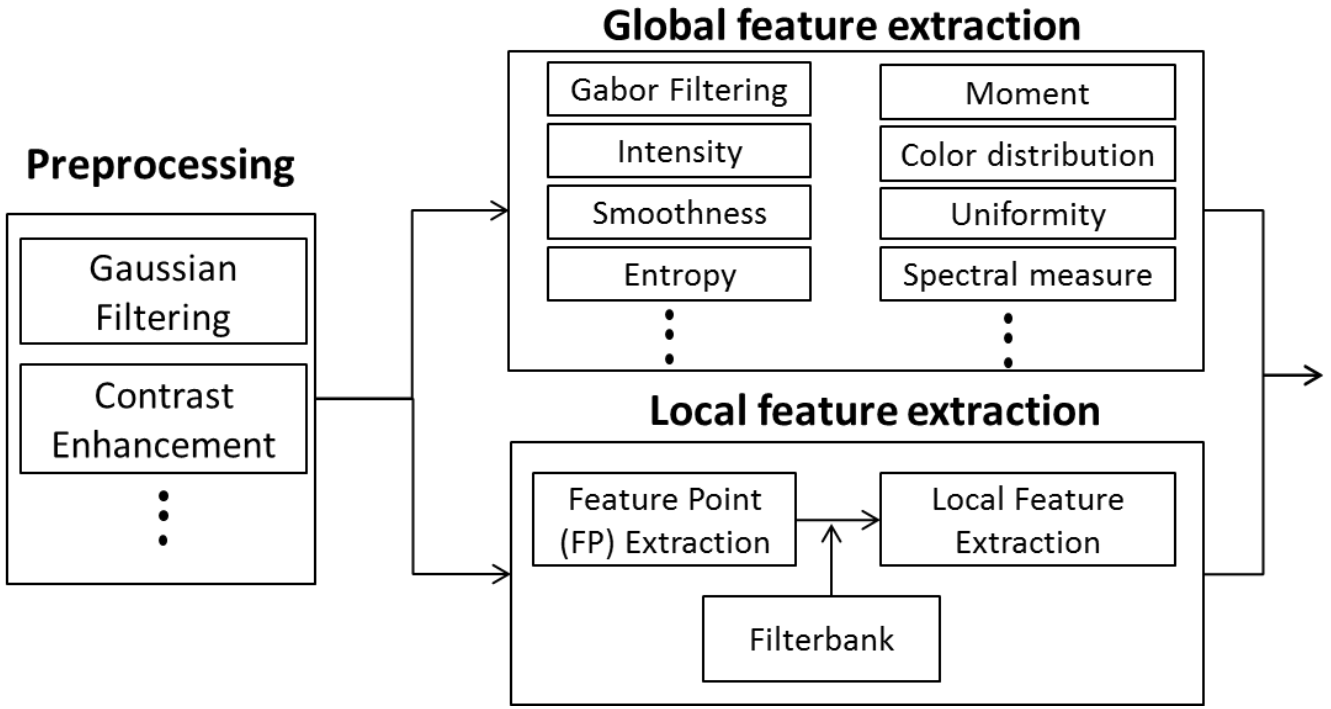
- AI to support safety assessment and surveillance
  - Drug toxicity curation through drug labeling
  - NAMs
  - Adverse reaction recognition and safety change tracking in drug labeling

Stage one of AI

- AI/ML require numerical features as the input.
  - Simplification: y = f(x)
- Custom codes to generate features, often called feature engineering
  - For images: edges, areas, size, texture, various transforms, …
  - For chemical structures: molecular descriptors, fingerprints
  - For text: word frequencies, co-occurrences, word types
- Common supervised AI/ML methods
  - Linear/Logistic Regression, KNN, Decision Trees, Random Forests, SVMs, neural network

An example - beetle species identification for food safety testing

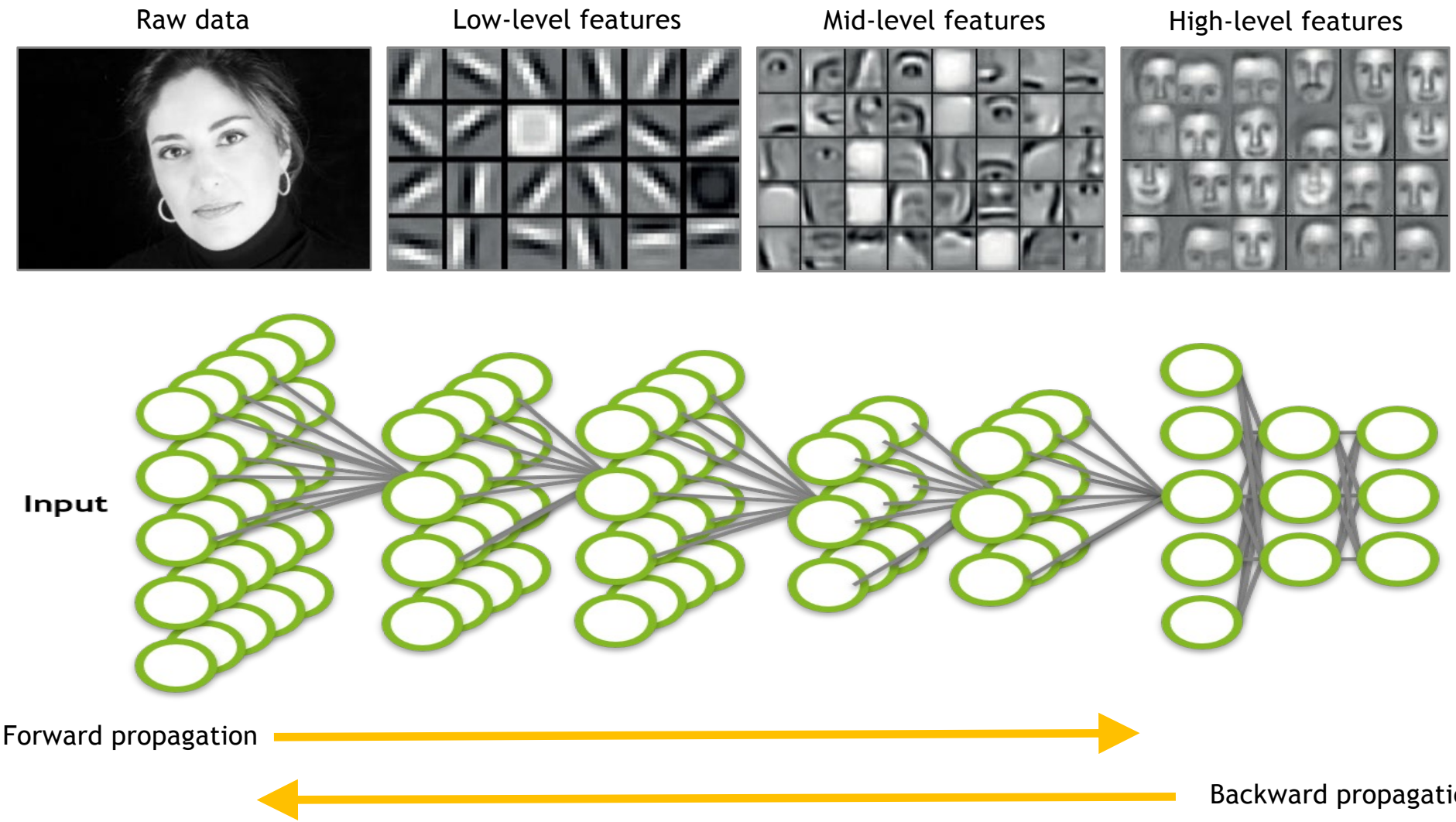
Feature extraction



ANN vs SVM

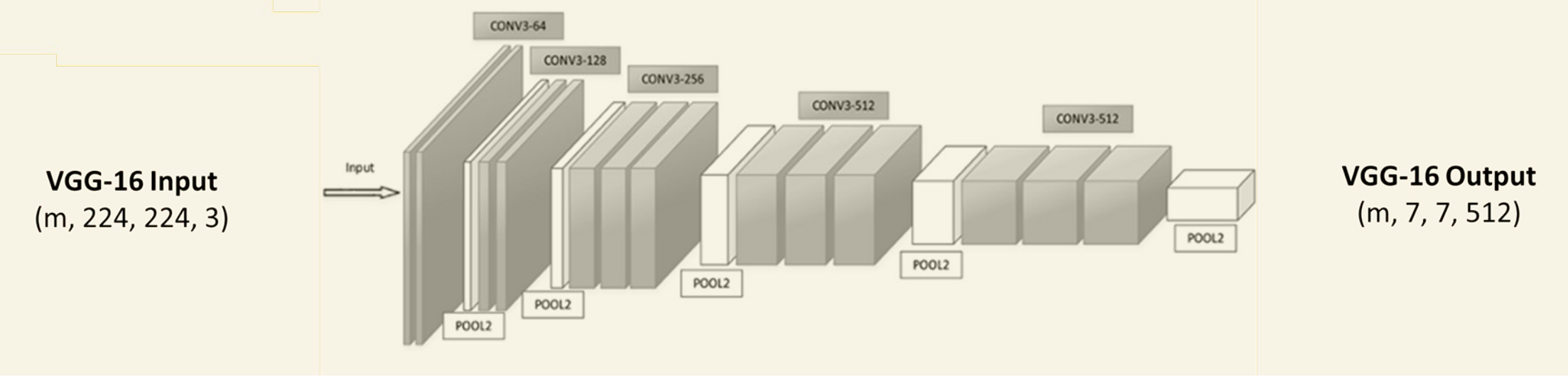
Park et al, PLOS One 11(6): e0157940 (2016)
Bisgin et al, Scientific Reports 8, 6532 (2018)

Stage two: Deep Learning



Advantages of transfer learning

- Bypass custom feature extraction methods which may have limited applicability
- Transfer learning (pretrained network, VGG-16) to bypass big data requirement
- Achieve better results for random image orientations: overall accuracy 79% vs 70%
- Better handling of magnification scale, image contrast and brightness



- However
  - Require a lot more computing resources such as HPC or GPU
  - Training takes much longer time; Some constraints for Deployment

Wu, et al, Computers and Electronics in Agriculture 166, 105002 (2019)

Stage three: foundational models

nature medicine

Article

https://doi.org/10.1038/s41591-024-02856-4

A visual-language foundation model for computational pathology

Received: 2 August 2023

Accepted: 5 February 2024

Published online: 19 March 2024

Check for updates

Ming Y. Lu[1,2,3,4,5,11], Bowen Chen[1,2,11], Drew F. K. Williamson[1,2,3,11], Richard J. Chen[1,2,3,4,6], Ivy Liang[1,7], Tong Ding[1,7], Guillaume Jaume[1,2,3,4], Igor Odintsov[1], Long Phi Le[2], Georg Gerber[1], Anil V. Parwani[8], Andrew Zhang[1,2,3,4,9] & Faisal Mahmood[1,2,3,4,10]

The accelerated adoption of digital pathology and advances in deep learning have enabled the development of robust models for various pathology

Advantages illustrated by CONCH

- Shortcomings of conventional models
  - Training is often difficult due to label scarcity in the medical domain
  - Model's usage is limited by the specific task and disease for which it is trained.
  - Most models leverage only image data, a stark contrast to how humans teach each other and reason about histopathologic entities.
- CONCH stands for CONtrastive learning from Captions for Histopathology
  - A visual-language foundation model
  - Trained over 1.17 million image–caption pairs through task-agnostic pretraining.
  - Evaluated on a suite of 14 diverse benchmarks, achieving state-of-the-art performance on histology image classification, segmentation, captioning, and text-to-image and image-to-text retrieval.
  - Using eight 80-GB NVIDIA A100 GPUs

MoLFormer-XL: A Foundation Model for Chemistry

Category	Key Details
Core Concept	A Transformer-based model pre-trained on **1.1 billion** unlabeled SMILES strings to learn a universal "chemical grammar".
Innovation	Uses **Linear Attention** for scaling and **Rotary Positional Embeddings (RoPE)** to capture 3D spatial relationships from 1D text.
Performance	Outperforms existing baselines on over 15 benchmarks, including solubility, toxicity, and quantum property prediction.
Applications	Accelerates drug discovery through rapid molecular screening, virtual screening, and de novo molecular generation.

Ross, J. et al. (2022). Nature Machine Intelligence, 4(12).
Ross, J. et al. (2024). arXiv:2405.04912 (GP-MoLFormer).

Part 2: AI to support safety assessment and surveillance

Drug-Induced Toxicity Classification Using FDA Labeling Documents

- Manual annotation of the labeling documents for drug toxicity classification is expensive:
  - It is a labor-intensive and expert-dependent process.
  - The classification needs to revised frequently in accordance with the revised labeling information.
  - New drugs need to be routinely reviewed.

DILIrank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans

Minjun Chen[1], Ayako Suzuki[2], Shraddha Thakkar[1], Ke Yu[1], Chuchu Hu[1] and Weida Tong[1]

[1] Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, USA
[2] Department of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR, USA

Generation of a drug-induced renal injury list to facilitate the development of new approach methodologies for nephrotoxicity

Skylar Connor[1], Ting Li[1], Yanyan Qu[1], Ruth A Roberts[2,3], Weida Tong[1,*]

[1] National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR 72079, USA
[2] ApconiX, Alderley Park, Alderley Edge SK10 4TG, UK
[3] University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

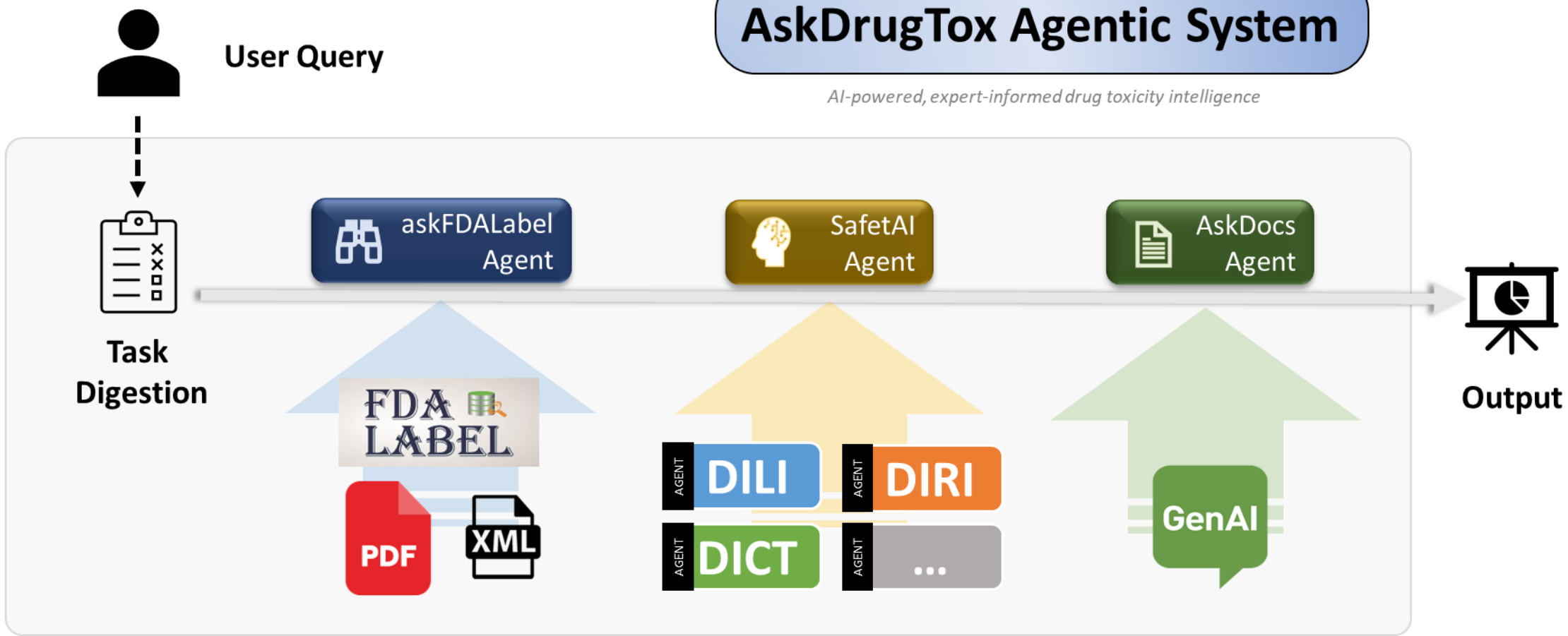
DICTrank: The largest reference list of 1318 human drugs ranked by risk of drug-induced cardiotoxicity using FDA labeling

Yanyan Qu[1,2], Ting Li[1], Zhichao Liu[1,†], Dongying Li[1,*], Weida Tong[1,*]

[1] National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, USA
[2] University of Arkansas at Little Rock and University of Arkansas for Medical Sciences Joint Bioinformatics Program, Little Rock, AR, USA

AskDrugTox Agentic System

AI-powered, expert-informed drug toxicity intelligence



- *Identify **tox-info** from trusted sources*
- *Assess risks in **similar products***
- *Keep **up-to-date** labeling data*

- *Specialized **agents** for each toxicity type*
- *Preserve **expert knowledge***
- *Fully **automated***

- *Deliver **reviewer-friendly** outputs*
- *Enable **Interactive** deep-dives*
- *Support Seamless **API integration***

AI Agent for DILI Classification

AI Agent Results

Human Results

	Predicted Positive	Predicted Negative
DILI Positive	149	1
DILI Negative	7	49

Of the 7 initially identified 'false positives', 6 were reclassified as true positives through manual verification of their new labeling data

DILIAgent can rapidly produce and update drug classification for DILI with sufficient accuracy

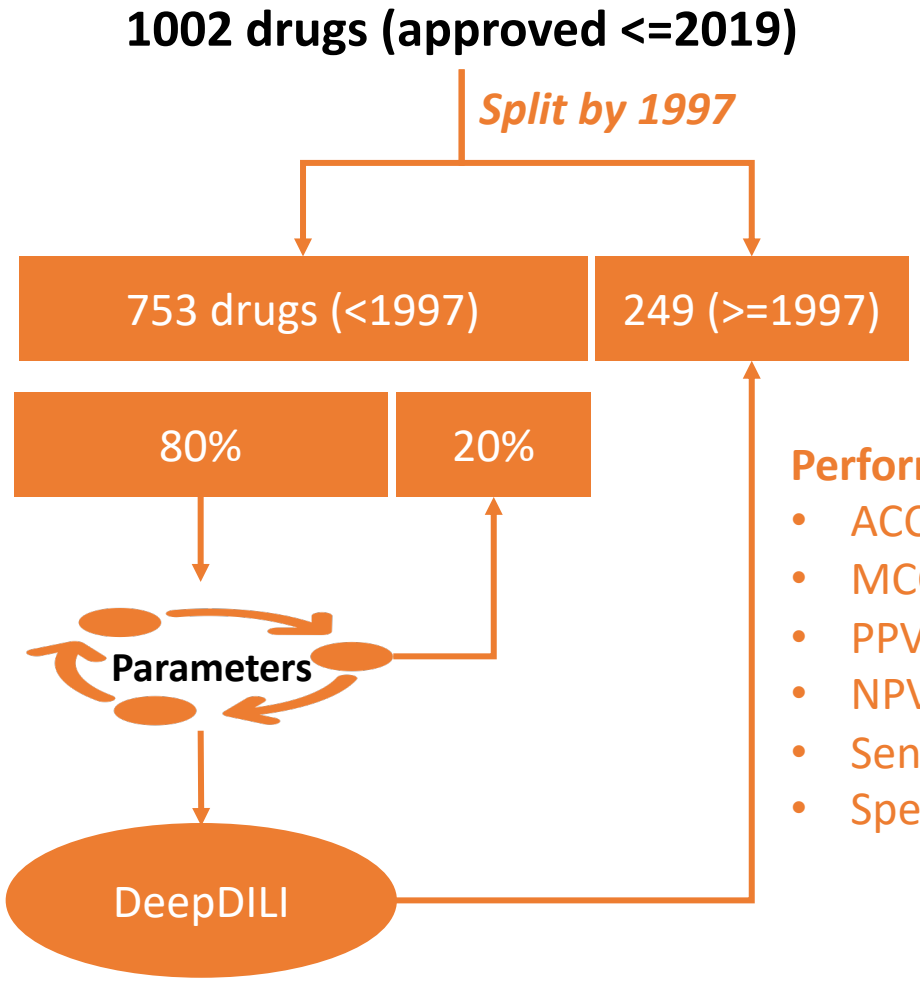
Minjun Chen … Weida Tong, Drug Discovery Today, 16(15-16):697-703, 2011

AI Agentic System for Drug Classification

- Save time and improve consistency
- Readily annotate new drugs and update for existing drugs
- Extend to other safety endpoints (e.g., cardiotoxicity, nephrotoxicity, etc)

- Support the development of New Approach Methods (NAMs) including in silico AI models (see examples in next slides)

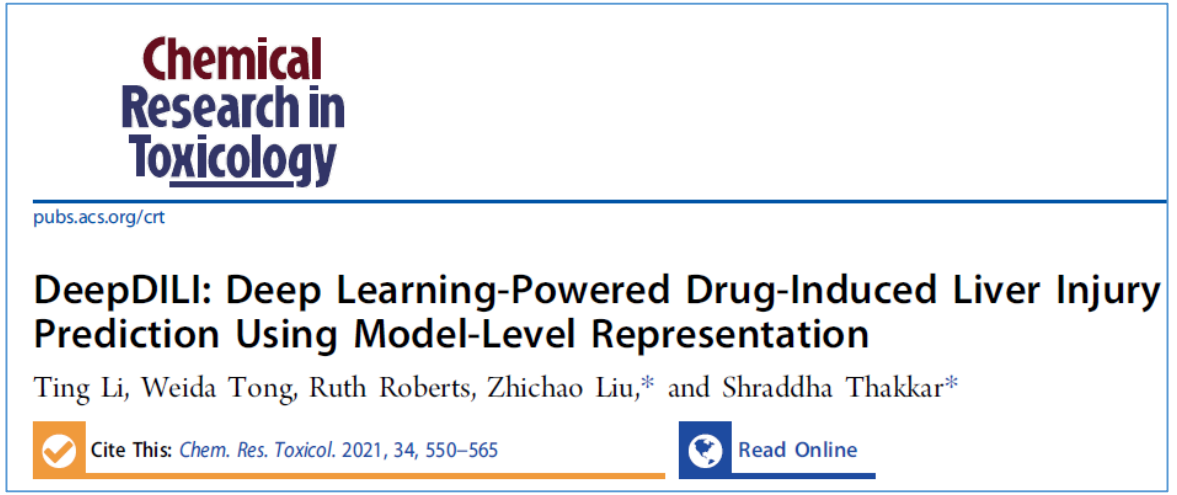
DeepDILI: an AI Model for DILI Assessment



- DeepDILI was developed with a **novel DL approach** that integrates multiple ML models in a deep learning architecture
  - A binary classification: DILI positive or negative
  - ~70% accuracy based on **drugs approved before 1997 to predict drugs approved after 1997**

Performance:

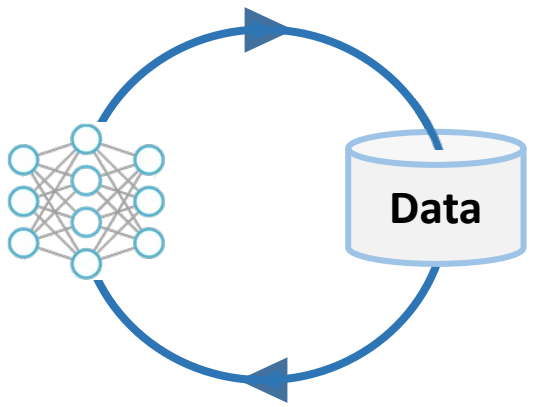
- ACC: 0.69
- MCC: 0.331
- PPV: 0.71
- NPV: 0.64
- Sens: 0.805
- Spec: 0.510



Li et al, CRT, 2021

Is an AI model improved over additional data?

Common Perception:
more data and a better model!



Adaptability behavior of DeepDILI in a real-world setting

Drugs approved before 1997 → Model1996

New drugs:

Model1996	
	1997-1998 (+53)
Model1998	
	1999-2001 (+44)
Model2001	
	2002-2004 (+46)
Model2004	
	2005-2007 (+45)
Model2007	
	2007-2019 (+61)
Model2019	

Questions:

1. Has the performance been improved?
2. Has the context-of-use been altered ?

Li et al, CRT, 2021

Example #3: Labeling AE Detection and Comparison

Why track AEs in labeling updates?

- Pre-approval clinical trials may not reveal all possible adverse reactions.
- Real-world usage after a drug's approval comprises larger and more diverse patient populations, relative to those in the pre-approval clinical trials.
- Labeling is then updated to include the previously undetected adverse reactions to provide accurate, up-to-date information on drug safety.
- Research into safety-related labeling changes including the frequency and timing and involved labeling sections is important to inform post-market surveillance of similar drug products.

Neyarapally, G.A., Wu, L., Xu, J. et al. Description and Validation of a Novel AI Tool, LabelComp, for the Identification of Adverse Event Changes in FDA Labeling. Drug Saf 47, 1265–1274 (2024).

RxBERT for Recognition and Extraction of Adverse Events from Drug Product Labeling

- RxBERT is a BERT-based model trained on drug product labeling, and evaluated its performance on named entity recognition of AEs
  - Training dataset included 44,900 drug product labeling
  - Custom regulatory language vocabulary ~28,000 unique terms

- Datasets: Text Analysis Conference (TAC) 2017 and FDA Adverse Drug Event Evaluation (ADE Eval 2021)

Wu L, Gray M, Dang O, Xu J, Fang H, Tong W. RxBERT: Enhancing drug labeling text mining and analysis with AI language modeling. Exp Biol Med (Maywood). 2023 Nov;248(21):1937-1943.

RxBERT for Recognition and Extraction of Adverse Events from Drug Product Labeling

Adverse reaction	Precision	Recall	*F*1
RxBERT	**89.3**	**88.5**	**88.9**
TAC challenge—max	86.4	86.9	85.2
TAC challenge—median	78.6	70.8	72.7
TAC challenge—min	42.1	13.4	20.3

Exact mention match—weighted	Precision	Recall	*F*1
RxBERT—ADE Eval	90.8	84.3	87.4
**ADE Eval challenge—max*	92	86	89
**ADE Eval challenge—median*	87	75	80
**ADE Eval challenge—min*	75	19	31

*The result from original ADE Eval challenge only contains two digits.

Wu L, Gray M, Dang O, Xu J, Fang H, Tong W. RxBERT: Enhancing drug labeling text mining and analysis with AI language modeling. Exp Biol Med (Maywood). 2023 Nov;248(21):1937-1943.

Two styles of visualization to show differences

(a)

Sub-section differences

Summary of Changes. Please seclect a section for details:

Section 5.2 Section 5.3 Section 5.6 Section 6.1 Section 6.2 Section 7.1

Current Section: Section 6.2

Old Labeling Section:

6.2 Postmarketing Experience The following adverse reactions have been identified during post-approval use of ABILIFY.

Because these reactions are reported voluntarily from a population of uncertain size, it is not always possible to establish a causal relationship to drug exposure: occurrences of allergic reaction (anaphylactic reaction, angioedema, laryngospasm, pruritus/urticaria, or oropharyngeal spasm), pathological gambling, hiccups and blood glucose fluctuation.

New Labeling Section:

6.2 Postmarketing Experience The following adverse reactions have been identified during post-approval use of ABILIFY.

Because these reactions are reported voluntarily from a population of uncertain size, it is not always possible to establish a causal relationship to drug exposure: occurrences of allergic reaction (anaphylactic reaction, angioedema, laryngospasm, pruritus/urticaria, or oropharyngeal spasm), pathological gambling, hiccups, blood glucose fluctuation, oculogyric crisis, and drug reaction with eosinophilia and systemic symptoms (dress).

Two new AEs are highlighted

(b)

Sentence Level difference Table: Reset Table

Copy CSV

Search:

	Section	Type	Sent_MedDRA
	6.2	filter	filter
126	6.2	Old	- Because these reactions are reported voluntarily from a population of uncertain size, it is not always possible to establish a causal relationship to drug exposure: occurrences of *allergic reaction* (*anaphylactic reaction*, *angioedema*, *laryngospasm*, *pruritus/urticaria*, or *oropharyngeal spasm*), *pathological gambling*, *hiccups* and *blood glucose fluctuation*.
127	6.2	New	+ Because these reactions are reported voluntarily from a population of uncertain size, it is not always possible to establish a causal relationship to drug exposure: occurrences of *allergic reaction* (*anaphylactic reaction*, *angioedema*, *laryngospasm*, *pruritus/urticaria*, or *oropharyngeal spasm*), *pathological gambling*, *hiccups*, *blood glucose fluctuation*, *oculogyric crisis*, and *drug reaction with eosinophilia and systemic symptoms* (*DRESS*).

Previous 1 Next

Detected AEs were placed into three categories

Highlighted AE Terms in AE-related labeling sections Download

**AE-related labeling sections include the boxed warnings and section 4(contraindications), 5(warnings & precautions), 6(adverse reactions), and 7(drug interactions)*

- **Newly added:**

cardio-respiratory arrest Sec. 6.1 | respiratory tract infection Sec. 6.1 | upper respiratory tract infection Sec. 6.1 | dress Sec. 6.2

drug reaction with eosinophilia and systemic symptoms Sec. 6.2 | eosinophilia Sec. 6.2

- **Removed:**

schizophreniaa weight gain Sec. 5.3 | autistic disorder abilify Sec. 5.6 | autistic disorder abilify 0/73 Sec. 5.6

- **Belows are other existed AEs:**

***These AEs were observed in changed texts, but have already been reported elsewhere in the old version; this is **NOT** all labeling involved AE terms*

cerebrovascular adverse events Sec. 5.2 | dementia Sec. 5.2 | fatalities Sec. 5.2 | psychosis Sec. 5.2 | stroke Sec. 5.2 | transient ischemic attack Sec. 5.2

autistic disorder Sec. 5.6 | bipolar disorder Sec. 5.6 | change in body weight Sec. 5.6 | change in body weight in abilify-treated patients Sec. 5.6

change in fasting glucose in abilify-treated patients Sec. 5.6 | changes in blood lipid parameters Sec. 5.6 | changes in fasting glucose levels Sec. 5.6

changes in total cholesterol Sec. 5.6 | depressive disorder Sec. 5.6 | fasting ldl cholesterol Sec. 5.6 | fasting triglycerides Sec. 5.6 | high Sec. 5.6 | irritability Sec. 5.6

Summary

- AI has opened a whole new world for aspiring chemists like you to discover new drugs and predict drug safety.
  - AI has removed some technical barriers to adoption.
  - AI has greatly leveled the playing field.

- Albert Einstein: The true sign of intelligence is not knowledge but imagination.
  - AI already shows the sign of imagination.
  - How you use AI is up to your imagination.