

Single Day Event

Beyond Traditional Methods:
Implementing AI/ML in Clinical
Stats Programming

Naga Durga Suravajhula,
Premier Research



Disclaimer:



The opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of Premier research. Premier research does not guarantee the accuracy or reliability of the information provided herein.



Understanding of AI and ML



AI algorithms and machine learning models have a significant impact on the biotech industry.



AI and ML tools are being leveraged across the Clinical industry (time and Cost efficiency)



Artificial Intelligence



Machine Learning



Deep Learning

Artificial intelligence (AI)

is a popular branch of computer science that concerns with building “intelligent” smart machines capable of performing intelligent tasks.

It is a sophisticated set of smart technologies that are customized to seek out and learn from data or experience; and then perform tasks based on what they learned while improving through experience

Machine learning (ML)

is a branch of AI in which algorithms and statistical models are used to enable computers to perform specific tasks;

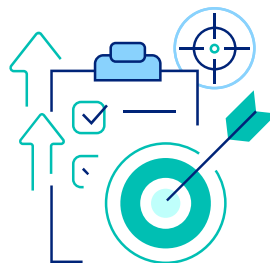
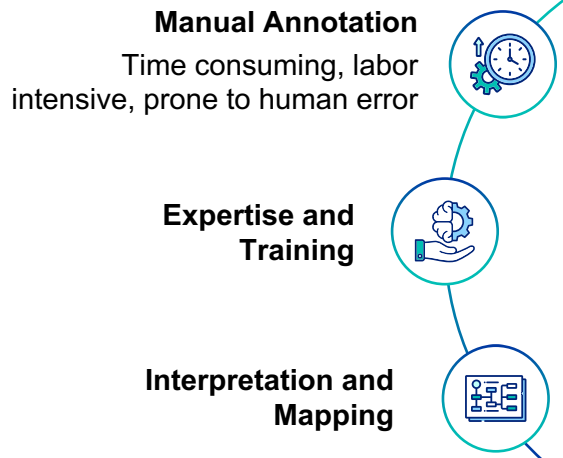
Over time the algorithm “learns” what patterns and information deliver the most useful results, then adapts future searches accordingly.

Based on the data you feed into machine learning algorithm and the training given to it, an output is delivered.

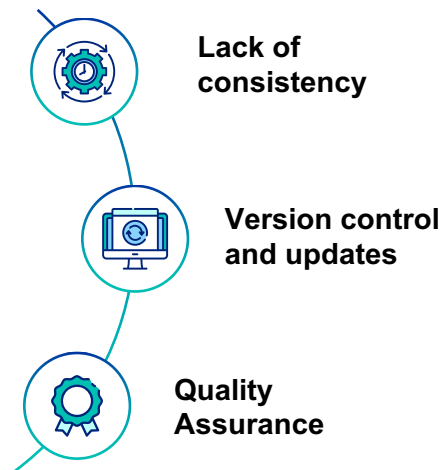


Challenges with Traditional Methods

Traditional methods of SDTM (Study Data Tabulation Model) CRF (Case Report Form) annotations and SDTM specification creations can involve several challenges and limitations



Challenges/ Limitations





Potential Benefits of AI/ML

Potential Benefits



Automated annotation and faster turnaround

- ☒ Improved efficiency
- ☒ Faster turnaround
- ☒ Reducing overall timelines
- ☒ Reduce manual effort
- ☐ Reduce resource requirement and cost



Improved consistency and Enhanced Accuracy

- ☐ Reduce variations
- ☐ Standardized and consistent
- ☒ Improved accuracy
- ☒ Reduce errors



Quality assurance

- ☐ Analyze changes
- ☐ Adapt the annotations to new SDTM standards
- ☐ Compliance with evolving standards



Continuous learning and Improvement

- ☐ Models can improve accuracy overtime
- ☐ Active learning



Practical Implementation of AI/ML

Input



CRF Metadata in Excel format

Folder	Form	Question Name	SDTM	Max Length	Type	Options
Screening	Demographics	Whatisthesubjectsyearofbirth	BRTHYR	4	TEXT BOX	null
Screening	Demographics	Whatisthesexofthesubject	SEX		BUTTONS	[[{sValue:MALE,sCodedValue:M},{sValue:FEMALE,sCodedValue:F}]]
Screening	Demographics	Whichofthefollowingracialdesignationsbestdescribesyou	RACE		SENTENCE OPTION	[[{sValue:AMERICAN INDIAN OR ALASKA NATIVE,sCodedValue:A}, {sValue:BLACK OR AFRICAN AMERICAN,sCodedValue:B}, {sValue:WHITE,sCodedValue:C}, {sValue:ASIAN,sCodedValue:D}, {sValue:HISPANIC OR LATINO,sCodedValue:E}]]
Screening	Demographics	WhatwasthedateofInformedConsent	DSSTDAT	17	DATE	{AllowOngoing:False,AllowFutureDates:False,AllowPastDates:True}
Screening	Demographics	Didthesubjectpreviouslyfailscreeninginthisstudy	PRVSBSF		BUTTONS	[[{sValue:YES,sCodedValue:Y},{sValue:NO,sCodedValue:N}]]
Screening	Demographics	IfYesPreviousSubjectIdentifier	PRSUBFID	7	TEXT BOX	null
Screening	Demographics	IWRSNumber	IWRSNUM	3	TEXT BOX	null
Screening	Demographics	eCRFVersion	CRFVER		BUTTONS	[[{sValue:3,sCodedValue:3}]]



Practical Implementation of AI/ML



ML recommendations

Demographics					DM		
Source					SDTM Target		
Item OID	Question	Control Type	Data Format	Data Type	SDTM Annotation	Variable	Label
BRTHYR	What is the subject's year of birth?	Text	0000	Char	BRTHDTC	DM-BRTHDTC	Date/Time of Birth
CRFVER	eCRF Version	Text		Char	BRTHDTC	DM-BRTHDTC	Date/Time of Birth
DSSTDAT	What was the date of Informed Consent?	DateTime	DD-MMM-YYYY	Char	RFICDTC	DM-RFICDTC	Date/Time of Informed Consent
IWRNUM	IWRS Number	Text	000	Char	BRTHDTC	DM-BRTHDTC	Date/Time of Birth
PRSUBFID	If Yes, Previous Subject Identifier	Text		Char	SUBJID	DM-SUBJID	Subject Identifier for the Study
PRVSUBSF	Did the subject previously fail screening in this study?	Text		Char	SUBJID	DM-SUBJID	Subject Identifier for the Study
RACE	Which of the following racial designations best describes you?	Text		Char	RACE	DM-RACE	Race
SEX	What is the sex of the subject?	Text		Char	SEX	DM-SEX	Sex



Practical Implementation of AI/ML



ML recommendations (Edited)

7

Demographics

DEMOG

DM

Source					SDTM Target		
Item ID	Question	Control Type	Data Format	Data Type	SDTM Annotation	Variable	Label
BRTHYR	What is the subject's year of birth?	Text	0000	Char	BRTHDTC	DM-BRTHDTC	Date/Time of Birth
CRFVER	eCRF Version	Text		Char	SUPPDM.QVAL WHEN QNAM = "CRFVER"	DM-BRTHDTC	Date/Time of Birth
DSSTDAT	What was the date of Informed Consent?	Date/Time	DD-MMM-YYYY	Char	RFICDTC	DM-QMDTC	Date/Time of Collection
IWRNSNUM	IWRS Number	Text	000	Char	SUPPDM.QVAL WHEN QNAM = "IWRNSNUM"	DM-BRTHDTC	Date/Time of Birth
PRSUBID	If Yes, Previous Subject Identifier	Text		Char	SUPPDM.QVAL WHEN QNAM = "PRSUBID"	DM-STUDYID	Study Identifier
PRYSUBSF	Did the subject previously fail screening in this study?	Text		Char	SUPPDM.QVAL WHEN QNAM = "PRYSUBSF"	DM-SUBJID	Subject Identifier for the Study
RACE	Which of the following racial designations best describes you?	Text		Char	RACE	DM-RACE	Race
SEX	What is the sex of the subject?	Text		Char	SEX	DM-SEX	Sex

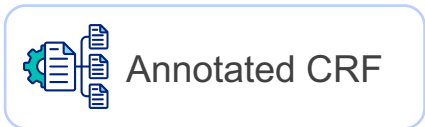
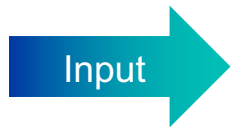
Validate

Download

Review/Approval



Practical Implementation of AI/ML



Form Items:

1	DEMOG.BRTHYR	What is the subject's year of birth? BRTHDTC		
2	DEMOG.CRFVER	eCRF Version SUPPDM.QVAL WHEN QNAM = "CRFVER"		
3	DEMOG.DSSTDAT	What was the date of Informed Consent? RFICDTC		
4	DEMOG.IWRGNUM	IWRG Number SUPPDM.QVAL WHEN QNAM = "IWRGNUM"		
5	DEMOG.PRISUBID	If Yes, Previous Subject Identifier SUPPDM.QVAL WHEN QNAM = "PRISUBID"		
6	DEMOG.PRVSUBSF	Did the subject previously fail screening in this study? SUPPDM.QVAL WHEN QNAM = "PRISOFI"		
7	DEMOG.RACE	Which of the following racial designations best describes you? RACE		
8	DEMOG.SEX	What is the sex of the subject? SEX		



Practical Implementation of AI/ML

Input



SDTM mapping specification (view in tool)

2	DM				Demographics		One record per subject				
Total Variables : 32											
Variable Name	Variable Label	Type	Controlled Terms	Core	Source Dataset	Source Variable	Mapping Algorithm/Comment	Origin	Active/Inactive	Edit	
STUDYID	Study Identifier	Char		Req	raw.DEMOG	DEMOG.PRSUBFID	STUDYID= ██████████				
DOMAIN	Domain Abbreviation	Char		Req			DOMAIN="DM";				
USUBJID	Unique Subject Identifier	Char		Req			USUBJID=cats('-',STUDYID,'01',SUBNUM);				
SUBJID	Subject Identifier for the Study	Char		Req	raw.DEMOG	DEMOG.PRVSUBSF	SUBJID = CATX('-', '01',SUBNUM);				
RFSTDTC	Subject Reference Start Date/Time	Char		Exp			by SUBNUM; if first.SUBNUM and TESTDATE ne . and TESTSTIM ne . then RFSTDTC= put(TESTDATE,js8601da.) "T" put(TESTSTIM,tod8.); else if first.SUBNUM and TESTDATE ne . and TESTSTIM eq . then RFSTDTC= put(TESTDATE,js8601da.);				
RFENDTC	Subject Reference End Date/Time	Char		Exp			by SUBNUM; if first.SUBNUM and TESTDATE ne . and TESTSTIM ne . then RFENDTC= put(TESTDATE,js8601da.) "T" put(TESTSTIM,tod8.); else if first.SUBNUM and TESTDATE ne . and TESTSTIM eq . then RFENDTC= put(TESTDATE,js8601da.);				
RFXSTDTC	Date/Time of First Study Treatment	Char		Exp			RFXSTDTC=RFSTDTC;				
RFXENDTC	Date/Time of Last Study Treatment	Char		Exp			RFXENDTC=RFENDTC;				



Practical Implementation of AI/ML

Input

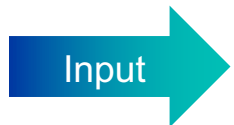


SDTM mapping specification (in excel format)

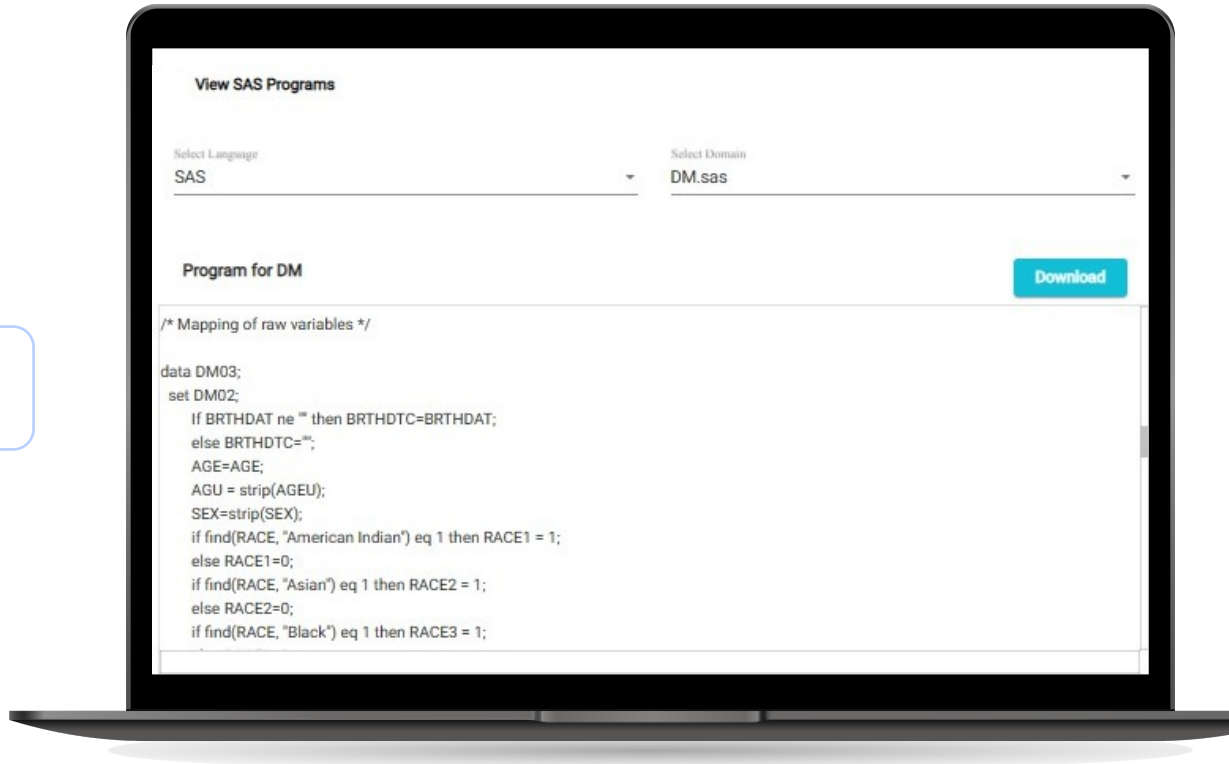
Order	Variable	label	type	codelist	core	Source_dataset	Source_Automation	source_variable	Mapping algorithm/comment
1	STUDYID	Study Identifier	Char		Req				STUDYID="XXXXXXXXXX"
2	DOMAIN	DomainAbbreviation	Char		Req				DOMAIN="DM";
3	USUBJID	Unique Subject Identifier	Char		Req	raw.DM	raw.DM.set	SITENUM, SUBNUM	USUBJID=cats('-',STUDYID,"01",SUBNUM);
4	SUBJID	Subject Identifierfor the Study	Char		Req	raw.DM	raw.DM.set	SUBID	SUBJID = CATX("-", "01",SUBNUM);
5	RFSTDTC	Subject Reference Start Date/Time	Char		Exp	raw.DM, raw.IP	raw.DM.merge, raw.IP.merge	IP.SUBNUM, IP.TESTDATE, IP.TESTSTIM	by SUBNUM; if first.SUBNUM and TESTDATE ne . and TESTSTIM ne . then RFSTDTC= put(TESTDATE,is8601da.) "T" put(TESTSTIM,tod8.); else if first.SUBNUM and TESTDATE ne . and TESTSTIM eq . then RFSTDTC= put(TESTDATE,is8601da.);
6	RFENDTC	Subject Reference End Date/Time	Char		Exp	raw.DM, raw.IP	raw.DM.merge, raw.IP.merge	IP.SUBNUM, IP.TESTDATE, IP.TESTSTIM	by SUBNUM; if first.SUBNUM and TESTDATE ne . and TESTSTIM ne . then RFENDTC= put(TESTDATE,is8601da.) "T" put(TESTSTIM,tod8.); else if first.SUBNUM and TESTDATE ne . and TESTSTIM eq . then RFENDTC= put(TESTDATE,is8601da.);
7	RFXSTDTC	Date/Time of FirstStudy Treatment	Char		Exp	SDTM.DM	SDTM.DM.set	RFSTDTC	RFXSTDTC=RFSTDTC;
8	RFXENDTC	Date/Time of Last Study Treatment	Char		Exp	SDTM.DM	SDTM.DM.set	RFENDTC	RFXENDTC=RFENDTC;



Practical Implementation of AI/ML



Auto-generated
SAS program





Potential Areas for Innovation and Advancement(AI/ML in Clinical Trials)



**Study Design –
eProtocol design**



**Study setup –
eCRF design**



Trial Management



**Data management –
smart queries, query
management, coding**



Data analysis



**Regulatory submission –
eTMF, CSR automation**

Thank You!



**The Global Healthcare
Data Science Community**

Contact Channels

📞 UK +44 1843 609600
✉ office@phuse.global
🌐 phuse.global

Social Media

🐦 [@phusetwitta](https://twitter.com/phusetwitta)
📘 [/phusebook](https://www.facebook.com/phusebook)
▶ [/phusetube](https://www.youtube.com/channel/UCphusetube)
🌐 [/company/phuse](https://www.linkedin.com/company/phuse)