- NLP (Natural language processing) is branch of Artificial intelligence focused on developing system which is based on natural communication language.

- Texts, spoken words, recorded sound etc. can be treated as input for NLP based algorithm

- NLP is mainly used to extract structure or meaning from natural communication language.

- Machine learning-predictive models are used to classify the data based on target value.

- Logistic regression, Decision tree etc. are example of classification algorithm.
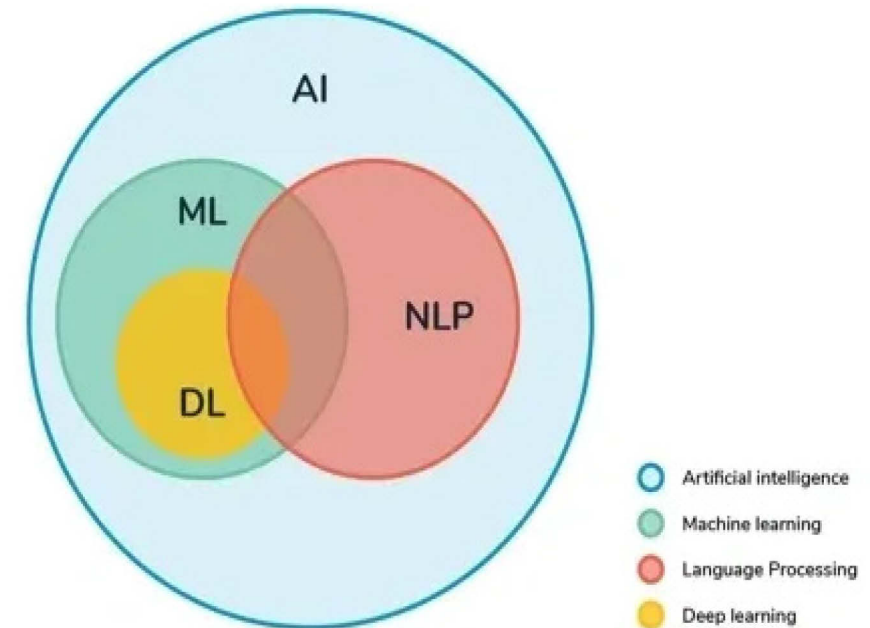
Detection of Offensive Language in Social Media Posts. - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Relationship-between-AI-ML-DL-and-NLP-7_fig8_343079524 [accessed 9 Sep, 2022]

Why NLP is special for SDLC

SDLC (Software development life cycle)

Analysis → Design → Development & Testing → Deployment → Maintenance

Note : This figure only represents steps irrespective of any methodology

All the phases in SDLC is meant for accomplishing some task, which is crucial for producing a quality software.

Each phases have different purpose, produce different result.

However, There is something which is very common in each phases in cycle.

Text Document- lot of data

To apply AI-backed algorithm, NLP is best suited.

Our focus today is software defect

Our work

- All closed defect is having data and attributes associated with it.

- We consider defect title, defect descriptions, steps (if any) and scope as data for creating algorithm.

- We consider NLP techniques to structure the unstructured data.

- We consider defect title, defect description and steps as predictor and try to predict the scope it belongs to.

- Similar defects from past can also be populated.

Defect- NLP Processing

Defect id	Defect Text	Affected Component
Defect 1	Defect Text	A
Defect 2	Defect Text	A
Defect 3	Defect Text	B
Defect 4	Defect Text	C
Defect 5	Defect Text	C
…….		
Defect n	Defect Text	A

NLP Processin →

Defect id	Defect Text	Affected Component
Defect 1	Feature vector	A
Defect 2	Feature vector	A
Defect 3	Feature vector	B
Defect 4	Feature vector	C
Defect 5	Feature vector	C
…….		
Defect n	Feature vector	A

Machine learning Possibilities

Defect id	Defect Text	Affected Component
Defect 1	Feature vector	A
Defect 2	Feature vector	A
Defect 3	Feature vector	B
Defect 4	Feature vector	C
Defect 5	Feature vector	C
…….		
Defect n	Feature vector	A

Predictive model for "Affected Component"

Similarity score between different defect vectors

Defect analytics use cases [AI-aided]

As we have enormous amount of data, what are possibilities , which can emerge to sharpen whole defect process.

Defect scope classification

- Newly raised defect can be automatically assigned to correct scope
- Thus, it will reduce the effort required to triage the defect.

Similar defect recommendation

- Once assignment is done, AI can help in extracting similar defect from past.
- This can reduce the resolving time of defect.
- Thus, it can reduce the duration of overall defect resolution time.

Defect trend extraction

- Defects trend can be captured from incoming defects.
- Forecasting for defect can be possible.

Value

Tester

Tester can easily understand the defect scope and hence assign accordingly.

Developer

Developer can search for similar defects fixed in past, hence get hint about the solution.

PO/Management

Get glimpse of pain points. Can get defect forecast and hence plan accordingly.

Customer

If we have efficient defect cycle, customer gets quick and satisfying product.

Any further questions?

Please contact:
kumar.rajesh@siemens.com