Course Code	Course Title	L	T	P	J	C
CSE4049	**Speech and Language Processing using Deep Learning**	**3**	**0**	**2**	**0**	**4**
Pre-requisite	**CSE3043**					**Syllabus version**
						v. 1.0

Course Objectives (COB):

1. Be competent with fundamental concepts for natural language processing and automatic speech recognition
2. To understand technologies involved in developing speech and language applications.
3. To demonstrate use of deep learning for building applications in speech and natural language processing

Expected Course Outcomes (CO):

1. Describe ways to represent speech and words
2. Demonstrate the working of sequence models for text
3. Adapt a dialogue system to a specific domain
4. Use signal processing techniques to analyze/represent the speech signal
5. Execute trials of speech/language systems

Module:1	**N-gram Language Models**	**5 hours**

Overview of NLP - Understanding Words – Corpora – Bag of Words - Text normalization - N-grams – Smoothing - Perplexity

Module:2	**Word Embedding**	**5 hours**

Lexical & word semantics – Words and vectors – Cosine similarity – Vector-Space models – TF-IDF – Word2vec – Bias and embedding – Evaluating vector models

Module:3	**Sequence Processing**	**7 hours**

Text Classification – Sentiment Analysis – Recurrent neural networks – Context in RNNs – Transformer networks – Machine translation – Encoder-decoder RNNs – Attention – Beam search – Evaluation of translation.

Module:4	**Dialogue Systems**	**5 hours**

IR based question answering system – Entity linking – Knowledge based Q&A – Language models for Q&A – Evaluation of systems – Chatbots – Human dialogue – Frame based dialogue – Dialogue state architecture – Evaluating dialogue systems.

Module:5	**Speech Production and Perception**	**5 hours**

Fundamentals of speech production – Short-Term Fourier representation of Speech – Functions of the ear – Perception of sound – Vocal tract model

Module:6	**Speech Signal Processing**	**7 hours**

Short-Time analysis of the signal – Energy – Zero crossing – Autocorrelation – Short time Fourier analysis - Spectrogram – Filter-banks – Cepstrum – Linear Predictive Coding – Mel-Frequency Cepstrum

Module:7	Automatic Speech Recognition	10 hours
Automatic Speech recognition formulation – HMM based speech recognition – Isolated word recognition – Large vocabulary continuous speech recognition – Deep learning for language modeling and automatic speech recognition – Evaluation metrics. Speaker recognition – Alexa/Google assistant based application development.		
Module:8	**Contemporary Issues**	**2 hours**
	Total Lecture hours: **45 hours**	

Text Book(s)	
1.	Speech and Language Processing, 3rd Ed., Daniel Jurafsky & James H. Martin, 2020.
2.	Theory and Applications of Digital Speech Processing, Lawrence R. Rabiner, Ronald W. Schafe, 1st Edn. Pearson, 2010.
Reference Books	
1.	Digital Speech Processing Using Matlab, E. S. Gopi, Springer, 2014
2	Voice Applications for Alexa and Google Assistant, Dustin Coates, Manning Publications, 2019.
3	Speech and Audio Processing A MATLAB -based Approach, Ian Vince, McLoughlin, Cambridge Press, 2016.
4.	Natural Language Processing with TensorFlow, Thushan Ganegedara, Packt, 2018
5.	An Introduction to Voice Computing in Python, Jim Schwoebel, NeuroLex, 2018
6.	Text Analytics with Python, Dipanjan Sarkar, Apress, 2019

Mode of Evaluation: Continuous Assessment Test (CAT) / Assignment / Quiz / Final Assessment Test (FAT) / Seminar

List of challenging Experiments – Indicative		
1	Text preprocessing – Tokenization, handling special chars, Stemming, Lemmatization	2 hours
2	Building N-gram language Model and testing its perplexity.	2 hours
3	TF-IDF model, document similarity	2 hours
4	Creating/using Word2vec and Glov2vec models and testing their performance	2 hours
5	LSTM for generating prose	2 hours
6	Machine translation from German to English/ or Indian language to English	2 hours
7	Creating a chatbot like a hostel help desk.	2 hours
8	Speech signal analysis- FFT, spectrogram, LPC, MFCC features	2 hours
9	HMM based Isolated word recognition	2 hours
10	Kaldi/ Sphinx Continous speech recognition	4 hours

11	Alexa speech enabled application development		4 hours
12	Google voice API based speech transcription.		4 hours
	Total Laboratory Hours		30 hours
Mode of Assessment: Continuous Assessment Test(CAT) & Final Assessment Test (FAT)			
Recommended by Board of studies	8-2-2021		
Approved by Academic Council	61st	Date	18-2-2021