Subject Description Form

Subject Code	EIE4105
Subject Title	Multimodal Human Computer Interaction Technology
Credit Value	3
Level	4
Pre-requisite	For 42477: EIE3103 Digital Signals and Systems or EIE3124 Fundamentals of Machine Intelligence For 42470: EIE3312 Linear Systems
Co-requisite/ Exclusion	Nil
Objectives	This course aims at providing students with the theories and applications of multimodal human-computer interaction (HCI) technologies. In particular, it enables students to understand how machine learning and deep learning can be applied to various HCI systems.
Intended Subject Learning Outcomes	**Upon completion of the subject, students will be able to:** Category A: Professional/academic knowledge and skills 1. Understand the capability and benefits of various HCI technologies. 2. Understand the theories of machine learning and deep learning. 3. Understand how machine learning and deep learning can be applied to various HCI systems. Category B: Attributes for all-roundedness 4. Understand the creative process when designing solutions to a problem.
Subject Synopsis/ Indicative Syllabus	1. HCI and Their Applications 1.1 Applications of HCI in daily life 1.2 Advantages of multimodal HCI 1.3 Trends in HCI technologies 1.4 Virtual reality, augmented reality, mixed reality, and metaverse 1.5 Real-life examples of HCI 2. Fundamental of Statistical Learning 2.1 Probability and random variables 2.2 Probability densities and distributions 2.3 Sampling distributions 2.4 Expectations and covariance 2.5 Bayes rule and Bayes decision theory 2.6 Curse of dimensionality 3. Machine Learning for HCI 3.1 Structure of pattern recognition systems. 3.2 Unsupervised Learning: principal component analysis; Eigenface, K-means clustering; Gaussian mixture models; hidden Markov models 3.3 Supervised Learning: linear discriminant analysis; support vector machines 3.4 Deep Learning: deep neural networks; backpropagation; gradient-based optimization; convolutional neural networks; representation learning; deep learning development platforms 3.5 Applications to handwriting recognition and face recognition. 4. Voice Computing 4.1 Voice computing: Interacting with computer through voice

<table>
<tr><td></td><td>4.2 Acoustic features
4.3 HMM and DNN for speech recognition
4.4 Language modelling
4.5 Speaker recognition: GMM-UBM, GMM-SVM, i-vectors, x-vectors, DNN speaker embedding, LDA, and PLDA
4.6 Applications of voice computing: voice search, spoken dialog systems, natural language processing, speech emotion recognition, speaker recognition, voice cloning.</td></tr>
<tr><td>Teaching/Learning Methodology</td><td>Lectures: The subject matters will be delivered through lectures. Students will be engaged in the lectures through Q&A, discussions, and specially designed classroom activities.

Tutorials: During tutorials, students will work on/discuss some chosen topics. This will help strengthen the knowledge taught in lectures.

Laboratory and assignments: During laboratory exercises, students will perform hands-on tasks to practice what they have learned. They will evaluate performance of systems and design solutions to problems. The assignments will help students to review the knowledge taught in class.

While lectures and tutorials will help to achieve the professional outcomes, the open-ended questions in laboratory exercises and assignments will provide the chance to students to exercise their creatively in problem solving.</td></tr>
<tr><td>Assessment Methods in Alignment with Intended Subject Learning Outcomes</td><td>
<table>
<tr><th rowspan="2">Specific Assessment Methods/Tasks</th><th rowspan="2">% Weighting</th><th colspan="4">Intended Subject Learning Outcomes to be Assessed (Please tick as appropriate)</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th></tr>
<tr><td>1. Continuous Assessment (total: 50%)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>• Homework and assignments</td><td>15%</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr>
<tr><td>• Tests and Quizzes</td><td>20%</td><td>✓</td><td>✓</td><td>✓</td><td></td></tr>
<tr><td>• Laboratory exercises</td><td>15%</td><td></td><td></td><td>✓</td><td>✓</td></tr>
<tr><td>2. Examination</td><td>50%</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr>
<tr><td>Total</td><td>100%</td><td colspan="4"></td></tr>
</table>
Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes:

Assignment, homework, and laboratory exercises will require students to apply what they have learnt to solve problems. There will be open-ended questions that allow students to exercise their creativity in making design.

Examination and tests: They assess students' achievement of the learning outcomes in a more formal manner.</td></tr>
<tr><td>Student Study Effort Expected</td><td>
<table>
<tr><td>Class contact (time-tabled):</td><td></td></tr>
<tr><td>• Lecture</td><td>24 Hours</td></tr>
<tr><td>• Tutorial/Laboratory/Practice Classes</td><td>15 Hours</td></tr>
<tr><td>Other student study effort:</td><td></td></tr>
<tr><td>• Lecture: preview/review of notes; homework/assignment; preparation for test/quizzes/examination</td><td>36 Hours</td></tr>
<tr><td>• Tutorial/Laboratory/Practice Classes: preview of materials, revision and/or reports writing</td><td>30 Hours</td></tr>
</table>
</td></tr>
</table>

	Total student study effort:	**105 Hours**
Reading List and References	**Reference Materials:** 1. M.W. Mak and J.T. Chien, *Machine Learning for Speaker Recognition*, Cambridge University Press, 2020. 2. I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press, 2016. 3. S.Y. Kung, M.W. Mak and S.H. Lin, *Biometric Authentication: A Machine Learning Approach*, Prentice Hall, 2005. 4. R. Haeb-Umbach, *et al*. "Speech Processing for Digital Home Assistants: Combining Signal Processing with Deep-learning Techniques", *IEEE Signal Processing Magazine*, Nov. 2019. 5. C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006. 6. S.J.D. Prince, *Computer Vision: Models, Learning, and Inference*, Cambridge University Press, 2012. 7. J.P. Thiran, F. Marques and H. Bourlard, *Multimodal Signal Processing, Theory and Applications for Human Computer Interaction*, Elsevier, 2010. 8. S. Greengard, *Virtual Reality*, MIT Press Essential Knowledge Series, 2019.	
Last Updated	June 2022	
Prepared by	Prof. M.W. Mak	