Syllabus

**ECE 5526 Speech Recognition**

**Description**

This course introduces students to the rapidly developing field of automatic speech recognition. Its content is divided into three parts.

* Part I deals with background material in the acoustic theory of speech production, acoustic-phonetics, and signal representation.
* Part II describes algorithmic aspects of speech recognition systems including pattern classification, search algorithms, stochastic modeling, and language modeling techniques.
* Part III compares and contrasts the various approaches to speech recognition, and describes advanced techniques used for acoustic-phonetic modeling, robust speech recognition, speaker adaptation, processing paralinguistic information, speech understanding, and multimodal processing.

Introduced concepts will be reinforced through practical exercises using Carnegie Melon University Sphinx system and MATLAB exercises. In addition e-WUW, proprietary Wake-Up-Word speech recognition technology, will be introduced and make use in projects.

**Organization**

There will be two 90 minute [lectures](http://ocw.mit.edu/OcwWeb/Electrical-Engineering-and-Computer-Science/6-345Automatic-Speech-RecognitionSpring2003/LectureNotes/) per week. To facilitate the coverage of a large quantity of material, copies of the lecture viewgraphs will be made available from <http://my.fit.edu/~vkepuska/ece5526/> .

There will be no final exam for the course. Instead there might be two in-class quizzes each counting approximately 15% towards the final grade.

There will be weekly [assignments](http://ocw.mit.edu/OcwWeb/Electrical-Engineering-and-Computer-Science/6-345Automatic-Speech-RecognitionSpring2003/Assignments/) consisting of both problems and computer work, so that students will be able to gain hands-on experience with the materials covered. Assignments must be turned in by the due date. Solutions will be provided along with the graded assignments. Each of the (~nine) assignments will count approximately 5% towards the final grade.

During the last quarter of the course, assignments will end, and students will work on a term project that will count approximately 25% towards the final grade. Projects will be chosen in consultation with staff members, and typically involve creating and evaluating a speech recognizer along a dimension of interest to the student. Tool kits of key recognizer components will be provided when necessary, so that minimal programming skills are necessary.

**Schedule**

Lecture: Two Sessions/Week

**Staff**

Lecturer:

Veton Këpuska

Olin Engineering Building

Office # 353

Tel. (321) 674-7183

E-mail: vkepuska@fit.edu

**Reference Material**

1. Rabiner & Juang. *Fundamentals of Speech Recognition.* Prentice-Hall, 1993.
2. Daniel Jurafsky & James H. Martin. Speech and language processing: An Introduction to Natural Lanugage Processing, Computational Lingustics, and Speech Recogntion, Pearson, Prentice Hall, 2009
3. Huang, Acero & Hon. *Spoken Language Processing.* Prentice-Hall, 2001.
4. Jelinek. Statistical Methods for Speech Recognition. MIT Press, 1997.
5. Duda, Hart & Stork. *Pattern Classification.* Wiley & Sons, 2001.
6. Stevens. *Acoustic Phonetics*, MIT Press, 1998.