Dictionary:
Spoken CALL Shared Task

This dictionary contains technical terms concerning the shared task
Mario Kunstek, Cem Freimoser, Dominik Jülg

Plain False Accepts (FA$_1$)

False accept of a response which is linguistically incorrect, but has the right meaning

Cross False Accepts (FA$_2$)

False accept of a response which is not even semantically correct

CSU-K System

System developed by Cooperative State University Karlsruhe (CSU-K)

DNN-HMM Model

Deep Neural Network, Hidden Markov Model. Acoustic models achieving good speech recognition results over gaussian mixture models (GMM-Hmm)

DBN Structure

Dynamic Bayesian Network. A generalization of HMM. Evolves over time. Umwandlung von durchgehende in diskrete Daten mit Genauigkeit delta.

Softmax Layer

Often used in the final layer of a neural network. Used for multi-class classification.

Forcing a DNN to output a probability. Converts number in a distribution so you obtain the probability

Bigram Model

Model for speech recognition. It is a special case of a N-gram

WER

Word error rate. A common metric for the performance of a speech recognition system

N-gram

A contiguous sequence of **n** items from a given sequence of text or speech. The items can be phonemes, syllables, letters, words or base pairs according to the application. Ähnlich einem Automat.

RASP System

Domain-independent robust parsing system for english

Linear Discriminant Analysis (LDA)

used in statistics, pattern recognition and machine learning to find a linear combination of features that characterizes or separates two or more classes of objects or events.

Maximum Likelihood Linear Transformation (MLLT)

computes a set of transformations that will reduce the mismatch between an initial model set and the adaptation data. More specifically MLLR is a model adaptation technique that estimates a set of linear transformations for the mean and variance parameters of a Gaussian mixture HMM system

Speaker Adaptive Training (SAT)

In acoustic modeling, speaker adaptive training (SAT) has been a long-standing

technique for the traditional Gaussian mixture models (GMMs). Acoustic models trained with SAT become independent of training speakers and generalize better to unseen testing speakers

Triphone model

In linguistics, a triphone is a sequence of three phonemes (one of the units of sound by which a word can be distinguished e.g. thin vs. din). Triphones are useful in models of natural language processing where they are used to establish the various contexts in which a phoneme can occur in a particular natural language

SRILM Tool

The SRI Language Modeling (SRILM) toolkit offers tools for building and applying statistical language models (LMs)

ASR

Automatic Speech Recognition

Language Model

A statistical **language model** is a probability distribution over sequences of words. Given such a sequence, say of length m, it assigns a probability to the whole sequence. When you look at the probability of a 3 word sequence you look at a 3-gram.

Sentence-Embedding Model

See Doc2Vec.

Word-Embedding Model

See Word2Vec. These are representation of words in a n-dimensional vector space so that semantically similar (e.g. "boat" — "ship") or semantically related (e.g. "boat" — "water") words come closer depending on the training method. Anders als bei einer discret representation (z.b [0000100000] werden bei Word-Embedding Model die

Vektoren aufgrund einer Verteilungsfunktion aufgebaut. $p(c \mid w)$ bildet die basis dieses Models. Es sagt auch wie Wahrscheinlich es ist, dass word w im kontext von c auftreten.

Dynamic Time Warping (DTW)

In time series analysis, dynamic time warping (DTW) is one of the algorithms for measuring similarity between two temporal sequences, which may vary in speed

Overfitting

Wenn man sein Netz zu gut trainiert dann ist das Pattern das er erkennen soll viel zu genau. D.h er kann mit störungen nicht mehr umgehen.

Logistic Regression

Logistic regression is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). The goal of logistic regression is to find the best fitting (yet biologically reasonable) model

Beam Search

In computer science, beam search is a heuristic search algorithm that explores a graph by expanding the most promising node in a limited set. Beam search is an optimization of best-first search that reduces its memory requirements.

LSTM

Long short term memory network.

Perplexity (ppl)

In information theory, perplexity is a measurement of how well a probability distribution or probability model predicts a sample. It may be used to compare probability models. A low perplexity indicates the probability distribution is good at predicting the sample.

Distributional similarity

Theorie aus der longisitc die ungefähr aussagt: "You should know a word by the company it keeps". Das bedeutet, dass die bedeutung eines Wortes dadurch bestimmt werden kann, indem man die wörter drum herum ansieht.

word2vec

Ist ein (bzw eine Menge von) Algorithmus der ein Textkonsomiert und dann ein word embedding Model erzeugt.

Hyperparameter

Parameter einer Verteilung die Gott-gegeben sind und daher nicht angepasst werden müssen. Z.b bei dem word2vec algo. Wäre solch ein parameter der Contextraum

Classification

Identifying to which category an object belongs to.

Regression

Predicting a continuous-valued attribute associated with an object.

Phoneme

Minimale einheit (ton) anhand derer zwei Wörter unterschieden werden.

Homophones

Wörter die nicht anhand ihrer aussprache, sondern nur anhand des kontext in welchen sie auftreten unterschieden werden können -> thier vs there

Prosody

Aussprache ändert bedeutung