What Is Claimed Is:

1. A speech processing method comprising:

receiving one or more speech signals, wherein each of the one or more speech signals comprises at least one dialogue uttered by a user;

extracting dialogues from the one or more speech signals;

identifying frequently uttered dialogues over a period of time, wherein the frequently uttered dialogues are a set of dialogues that are uttered by the user a number of times during the period of time more than other dialogues uttered by the user during the period of time;

determining that a size of a local language model and a local acoustic model will result in a sufficient quality threshold, wherein the size is based on, at least in part, the frequently uttered dialogues;

generating the local language model and the local acoustic model based on, at least in part, the frequently uttered dialogues when the size of the local language model and the local acoustic model will result in the sufficient quality threshold;

processing the one or more speech signals based on, at least in part, the local language model and the local acoustic model, wherein the extracting dialogues from the one or more speech signals further comprises implementing a remote language model and a remote acoustic model for extracting the dialogues from the one or more speech signals, and wherein the local language model and the local acoustic model are subsets of the remote language model and the remote acoustic model, respectively, wherein the local language model and the local acoustic model are executed on a user device and the remote language model and the remote acoustic model are executed remotely from the user device on at least one server, wherein the processing the one or more speech signals further comprises executing a local Automatic Speech Recognition (ASR) engine on a user device to process the one or more speech signals based on, at least in part, the local language model and the local acoustic model, wherein the local language model and the local acoustic model are trained for the local ASR engine by a remote ASR engine on the at least one server; and

receiving, by the local ASR engine, the local language model and the local acoustic model trained by the remote ASR engine.

2. The speech processing method according to claim 1, wherein the local language model and the local acoustic model are trained for the local ASR engine by the remote ASR engine offline and pushed back to the local ASR engine.

3. The speech processing method according to claim 1, further comprising categorizing the dialogues into different domains.

4. The speech processing method according to claim 3, further comprising, after categorizing the dialogues into different domains, ranking the categorized dialogues in each of the domains based on, at least in part, a frequency of utterance of each of the dialogues by the user over the period of time, such that the dialogues with a higher frequency of utterance are ranked higher as compared to the dialogues with a relatively lower frequency of utterance in the corresponding domain.

5. The speech processing method according to claim 4, wherein the identifying frequently uttered dialogues further comprises identifying a predefined number of dialogues with high rankings in each of the domains as the frequently uttered dialogues.

6. The speech processing method according to claim 5, wherein the predefined number of dialogues are determined based on, at least in part, one or more of a memory and a processor of a user device.

7. The speech processing method according to claim 4, wherein the domains include one or more of food, entertainment, sports, scheduling, sales inquiry, and automation commands.

8. The speech processing method according to any of claims 1-7, further comprising storing the local language model and the local acoustic model, wherein the local language model and the local acoustic model is based on, at least in part, frequently uttered dialogues over a period of time, wherein the frequently uttered dialogues are a set of dialogues that are uttered by the user a number of times

during the period of time more than other dialogues uttered by the user during the period of time.

9. The speech processing method according to any of claims 1-8, further comprising associating the local language model and the local acoustic model with a user profile of the user.

10. The speech processing method according to any of claims 1-9, wherein the storing the local language model and the local acoustic further comprises storing the local language model and the local acoustic model in a memory of a user device linked with the user profile of the user.

11. A speech processing system configured to perform the method according to any of the preceding claims..

12. A computer program product residing on a computer readable storage medium having a plurality of instructions stored thereon which, when executed across one or more processors, causes at least a portion of the one or more processors to perform the method according to any of claims 1-10.