SOFTWARE SYSTEM DEVELOPMENT

LEARNING A CODE LANGUAGE MODEL USING SELF-SUPERVISED LEARNING (SSL) FOR CODE

Team 22

NIKHIL SINGH	2024201067
RAJATH GADAGKAR	2024201035
ABHRADEEP DAS	2024202018
SANKET ADLAK	2024204005

ABOUT PROJECT

- **AIMED AT ENHANCING THE PRODUCTIVITY OF SOFTWARE DEVELOPERS**

- **CODE SUGGESTION TOOL**

- **SSL-BASED TRANSFORMER MODELS FOR PREDICTIONS**

- **PREDICTS THE NEXT SEGMENT OR TOKEN OF CODE**

- **REAL-TIME, CONTEXT-AWARE CODE COMPLETION**

- **WIDE DATASET COVERAGE**

- **FRONTEND INTERFACE FOR USER INTERACTION**

ASSUMPTIONS

- **OPEN-SOURCE DATA AVAILABILITY**

- **LOW-LATENCY PERFORMANCE**

- **USER FEEDBACK AS A LEARNING LOOP**

- **MODULAR DESIGN OF THE SYSTEM**

- **FUTURE LANGUAGE/FEATURE INTEGRATION**

TRANSFORMERS

- **REDUCES COMPUTE COSTS**

- **LIGHTENS CARBON FOOTPRINT**

- **SAVES THE TIME**

- **REDUCES MODEL TRAINING RESOURCES**

- **SUPPORTS FRAMEWORK INTEROPERABILITY**

TRANSFORMERS

- **TRANSFORMER-XL**

  - **ATTENTIVE LANGUAGE MODELS BEYOND A FIXED-LENGTH CONTEXT**

  - **TRANSFORMER-XL LEARNS DEPENDENCY THAT IS 80% LONGER THAN RNNS**

- **LSTM**

  - **DESIGNED TO LEARN AND PREDICT SEQUENTIAL DATA**

  - **EXCELS AT CAPTURING LONG-TERM DEPENDENCIES**

- **RECURRENT HIGHWAY NETWORK**

  - **EXTEND THE LSTM ARCHITECTURE TO ALLOW STEP-TO-STEP TRANSITION DEPTHS LARGER THAN ONE**

TECHNOLOGY STACK - BACKEND

- **FLASK - SERVES AUTOCOMPLETION REQUESTS**

- **PYTHON - PRIMARY LANGUAGE FOR BACKEND**

- **LABML - USED FOR LOGGING AND TRACKING THE PERFORMANCE**

- **LABML_NN - BUILDING AND TRAINING NEURAL NETWORKS**

- **TORCH - DEFINING, TRAINING, AND USING NEURAL NETWORK MODELS**

TECHNOLOGY STACK - FRONTEND

- **TYPESCRIPT**

    - **DEVELOP THE VS CODE EXTENSION**

    - **ASYNCHRONOUS COMMUNICATION WITH THE BACKEND**

- **NODE.JS**

    - **RUNTIME ENVIRONMENT**

    - **EVENT-DRIVEN COMMUNICATION**

CONTROL FLOW

DATA FLOW

PROJECT EVOLUTION

PHASE 1

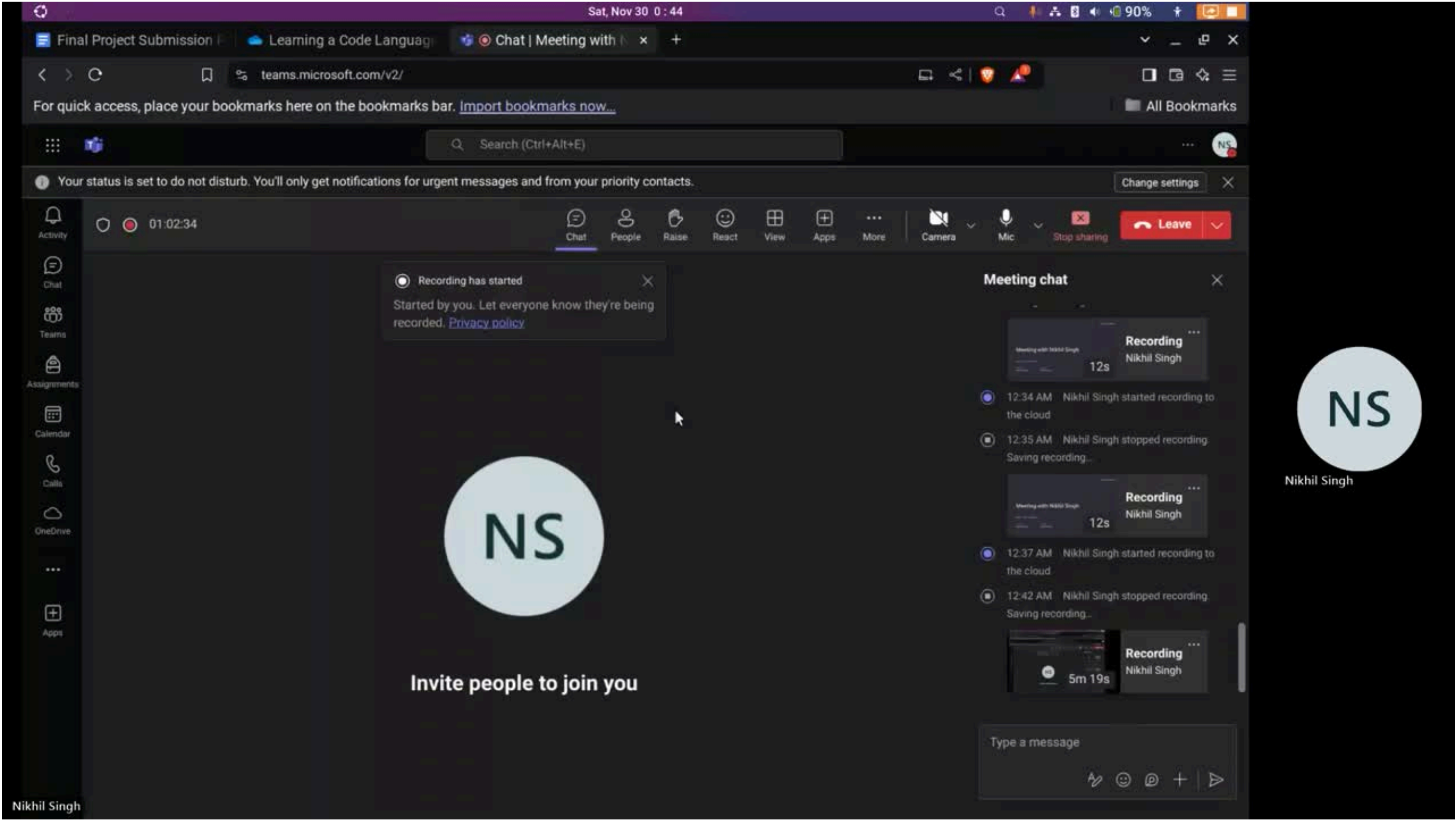- **RESEARCH, SYSTEM DESIGN, DATASET SELECTION, AND ARCHITECTURE.**

PHASE 2

- **DATASET PREPARATION, BASIC MODEL TRAINING, FRONTEND UI DESIGN.**

PHASE 3

- **VSCODE EXTENSION CREATION, REAL-TIME CODE AUTOCOMPLETION.**

VIDEO

THANK YOU!