PROJECT MEETING SUMMARY: BAYESIAN NETWORK FOR CARDIOVASCULAR DISEASE PREDICTION

PROJECT OVERVIEW

The team is developing a Bayesian network system to predict cardiovascular disease risk. The project uses a dataset of 7,000 individuals with 12 health attributes including smoking, alcohol intake, blood pressure, and cholesterol levels. The goal is to create a decision tree that clinicians and patients can use to assess disease risk.

KEY DISCUSSION POINTS

1. PROJECT SCOPE AND TIMELINE

- Work period: 7 weeks until finalizing for research publication
- Must complete core fundamentals before rushing to publication
- Focus on manageable workload with weekly progress milestones
- Scoped work to prevent burnout and ensure quality output

2. TECHNICAL COMPONENTS BREAKDOWN

Decision Tree Development

- Create manually by discussing team observations
- Also develop programmatically using algorithms
- Must establish logical causal relationships between factors
- Example: background factors (age, smoking) lead to symptoms, which inform diagnosis

Conditional Probability Tables (CPTs)

- Need to estimate probabilities for each variable relationship
- Will use Bayesian network libraries like PGMPY
- Research how existing projects estimate CPTs
- Consider CPT estimators for automated calculation

Visualization

- Create graphs showing decision tree structure
- Display how treatments affect probability outcomes
- Use Python libraries for implementation

3. DATA AND ATTRIBUTES DISCUSSION

Current Dataset Issues

- 12 attributes focus on background health factors and test results
- Missing clear symptom data that patients can self-report
- Examples of attributes: smoking, alcohol intake, age, blood pressure, cholesterol
- Attributes are not user-observable symptoms (chest pain, shortness of breath)

Symptoms vs. Test Results

- Background factors: age, smoking status, family history
- Test results: blood pressure (systolic/diastolic), cholesterol, glucose
- Patient symptoms: chest pain, shortness of breath (not in current dataset)

- Decision: focus on measurable background and test data; may research adding symptoms separately

4. WORKFLOW AND ROLE ASSIGNMENTS
- Someone handles data analysis, visualization, and research
- Someone focuses on decision tree and mathematical backing
- Someone works on conditional probability tables
- Ensuring equal contribution (33% each) and tangible deliverables per person

5. RESEARCH RECOMMENDATIONS

Bayesian Networks for Medical Diagnosis
- Investigate existing Bayesian network implementations for medical use
- Study how decision trees are constructed from symptom data
- Analyze how CPTs are estimated in clinical applications
- Look for comparable projects to validate approach feasibility

Tools and Methods
- Use AI tools (like Menace) to find research papers and generate decision trees
- Review existing decision trees from academic papers on cardiovascular disease
- May use programmatic generation of decision trees from data
- Compare manual vs. AI-generated decision trees for validation

6. TECHNICAL TOOLS AND SETUP
- Python with libraries: PGMPY, NumPy, Pandas
- Menace AI for research and document generation
- GitHub for version control and documentation
- Google Docs for shared project planning
- Anaconda or local development environment

7. DELIVERABLES FOR NEXT WEEK (by Monday)
- Updated proposal document with refined project scope
- Research articles identifying suitable decision tree structures
- Initial decision tree draft (manual sketch plus logical outline)
- Validation of whether identified research aligns with project goals
- If feasible, programming the decision tree from researched models

TEAM DYNAMICS
The team expressed enthusiasm about the project scope. They appreciated the discussion focusing the work to be manageable while maintaining quality. Clear role differentiation allows parallel work, enabling the conditional probability work to proceed once the decision tree is finalized.

NEXT STEPS
1. Redo proposal document before next meeting
2. Find and review research papers on Bayesian networks for cardiovascular disease
3. Create rough decision tree draft based on research findings
4. Validate approach with advisor before proceeding to implementation

5. Schedule next meeting for Monday (after scrum session)

TECHNICAL NOTES
- The system will not require users to input traditional symptoms since the dataset lacks them
- Instead, it will work with measurable health indicators and test results
- Treatment incorporation into the decision tree will create branches reflecting different outcomes
- Dynamic updates will allow real-time probability adjustments based on new variables