**Antes de modelar: Planejando e delimitando o escopo de um projeto**

Este capítulo aborda:

Definição de estratégias de planejamento eficazes para trabalhos de projeto de ML

Uso de métodos eficientes para avaliar possíveis soluções para um problema de ML.

Os dois maiores assassinos de projetos de ML no mundo não têm nada a ver com o que a maioria dos cientistas de dados imagina. Esses assassinos não estão relacionados a algoritmos, dados ou perspicácia técnica. Eles não têm absolutamente nada a ver com a plataforma que você está usando, nem com o mecanismo de processamento que otimizará um modelo. Os maiores motivos para os projetos não atenderem às necessidades de uma empresa estão nas etapas que levam a qualquer um desses aspectos técnicos: as fases de planejamento e definição do escopo de um projeto. Durante a maior parte da educação e treinamento que recebemos antes de trabalhar como DS em uma empresa, a ênfase é colocada fortemente na solução independente de problemas complexos. Isolar-se e focar em demonstrar habilidade demonstrável na compreensão da teoria e aplicação de algoritmos nos treina para ter a expectativa de que o trabalho que faremos na indústria seja individual. Dado um problema, descobrimos como resolvê-lo. A realidade da vida na capacidade de DS não poderia ir além da abordagem acadêmica de comprovar o conhecimento e a habilidade de alguém na resolução de problemas sozinho. Esta profissão é, na verdade, muito mais do que apenas algoritmos e acumular conhecimento sobre como utilizá-los. É um campo altamente colaborativo e orientado por pares; os projetos mais bem-sucedidos são construídos por equipes integradas de pessoas que trabalham juntas, comunicando-se durante todo o processo. Às vezes, esse isolamento é imposto pela cultura da empresa (isolar intencionalmente a equipe do resto da organização sob a intenção equivocada de “proteger” a equipe de solicitações aleatórias de projetos) e outras vezes é autoimposto. Este capítulo aborda por que essa mudança de paradigma em que as equipes de ML se concentram menos no como (algoritmos, tecnologia e trabalho independente) e mais no quê (comunicação e colaboração no que está sendo construído) pode contribuir para um projeto bem-sucedido. Essa mudança ajuda a reduzir o tempo de experimentação, concentrar a equipe na construção de uma solução que funcione para a empresa e planejar o trabalho do projeto em fases que incorpore o conhecimento das PMEs de equipes multifuncionais para ajudar a aumentar drasticamente as chances de um projeto bem-sucedido. O início desta jornada inclusiva, de reunir o maior número possível de pessoas para criar uma solução funcional que funcione para resolver um problema, está na fase de definição do escopo. Vamos justapor o fluxo de trabalho de uma equipe de ML que tem escopo e planejamento inadequados ou ausentes (figura 3.1) com um fluxo de trabalho que inclui escopo e planejamento adequados (figura 3.2).

Sem culpa alguma (a menos que queiramos culpar a equipe por não ser enérgica com a unidade de negócios para obter mais informações, o que não faremos), esses membros da equipe de ML fazem o possível para construir diversas soluções para resolver o vago requisitos lançados em seu caminho. Se tiverem sorte, acabarão com quatro MVPs e vários meses de esforço desperdiçados em três que nunca chegarão à produção (muito trabalho desperdiçado). Se tiverem muito azar, terão desperdiçado meses de esforço em nada que resolva o problema que a unidade de negócios deseja resolver. De qualquer forma, não há bons resultados. Com o escopo e o planejamento adequados mostrados na figura 3.2, o tempo gasto na construção de uma solução é reduzido consideravelmente. O maior motivo para essa mudança é que a equipe tem menos abordagens totais para validar (e todas têm um prazo de duas semanas), principalmente porque o feedback “antecipado e frequente” é recebido pelo cliente interno. Outra razão é que em cada fase do desenvolvimento de novos recursos, uma rápida reunião e demonstração da funcionalidade adicionada é realizada para testes de aceitação pelas PMEs. Para além da melhoria substancial da eficiência, o outro grande benefício desta metodologia de inclusão com o cliente interno é um aumento significativo da probabilidade de a solução final satisfazer as expectativas do negócio. Já se foi o risco extremo mostrado na figura 3.1: entregar demonstrações de múltiplas soluções após meses de trabalho, apenas para descobrir que todo o projeto precisa ser reiniciado do zero.

