

解决方案实践

讯飞星火代码大模型解决方案实践

文档版本 1.0
发布日期 2025-05-09



版权所有 © 华为技术有限公司 2025。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

目 录

1 方案概述..... 1

2 资源和成本规划..... 5

3 实施步骤..... 6

3.1 IFlyCode-VPC 管理后台（面向管理员） 6

3.2 iFlyCode 插件（以 IDEA 为例） 18

4 修订记录..... 29

1 方案概述

应用场景

客户痛点：

在软件开发过程中，企业和开发者常面临以下挑战：

- 低效的重复编码：大量时间耗费在样板代码、重复逻辑和基础架构搭建上，影响核心业务逻辑的开发效率。
- 知识壁垒与学习成本：新技术、新框架或遗留系统的理解和应用需要高昂的学习成本，拖慢项目进度。
- 代码质量与维护难题：人工编写的代码可能存在潜在缺陷、性能瓶颈或可读性问题，导致后期维护成本激增。
- 人才短缺与协作摩擦：团队技能差异导致代码风格不统一，新成员上手慢，跨团队协作效率低下。
- 安全与合规风险：手动编码可能引入安全隐患或违反行业合规要求，增加企业风险。
- 智能编程助手通过AI技术准确解决这些问题，赋能开发者更智能、更高效、更安全地编写代码。

【场景1】代码生成与智能补全

痛点：开发者需反复编写重复性代码（如CRUD接口、数据模型），占用创新时间。

解决方案：通过自然语言描述或上下文理解，自动生成高完整性代码，支持主流语言（Python、Java、Go等），大幅提升编码速度。

【场景2】实时错误检测与修复

痛点：调试耗时占开发周期的30%以上，且复杂错误依赖经验排查。

解决方案：动态分析代码逻辑，即时提示语法错误、类型冲突及潜在缺陷，并提供一键修复建议，降低调试成本。

【场景3】代码优化与重构

痛点：遗留代码性能差、可读性低，重构风险高、周期长。

解决方案：自动识别冗余代码，推荐算法优化方案，安全重构变量、函数及模块结构，保障代码健壮性。

【场景4】文档与测试自动化

痛点：文档缺失或过时，手动编写测试用例覆盖不全。

解决方案：根据代码生成标准化文档（API说明、函数注释），并自动创建单元测试模板，提升覆盖率至60%+。

【场景5】多语言与跨平台开发

痛点：跨技术栈开发需重复学习，适配成本高。

解决方案：智能翻译代码逻辑（如Python转C#），生成多平台兼容代码（Web/iOS/Android），加速全栈开发。

【场景6】安全与合规护航

痛点：手动编码易忽略漏洞（如SQL注入、敏感数据泄露）。

解决方案：深度扫描代码，标记OWASP Top 10风险，并提供合规性检查（GDPR、HIPAA），降低安全事件概率。

【场景7】团队协作与知识传承

痛点：新成员熟悉代码库需数周时间，协作效率低。

解决方案：通过AI解读代码逻辑，生成可视化流程图，并统一团队编码规范，缩短onboarding 周。

方案架构

业务架构

图 1-1 业务架构图

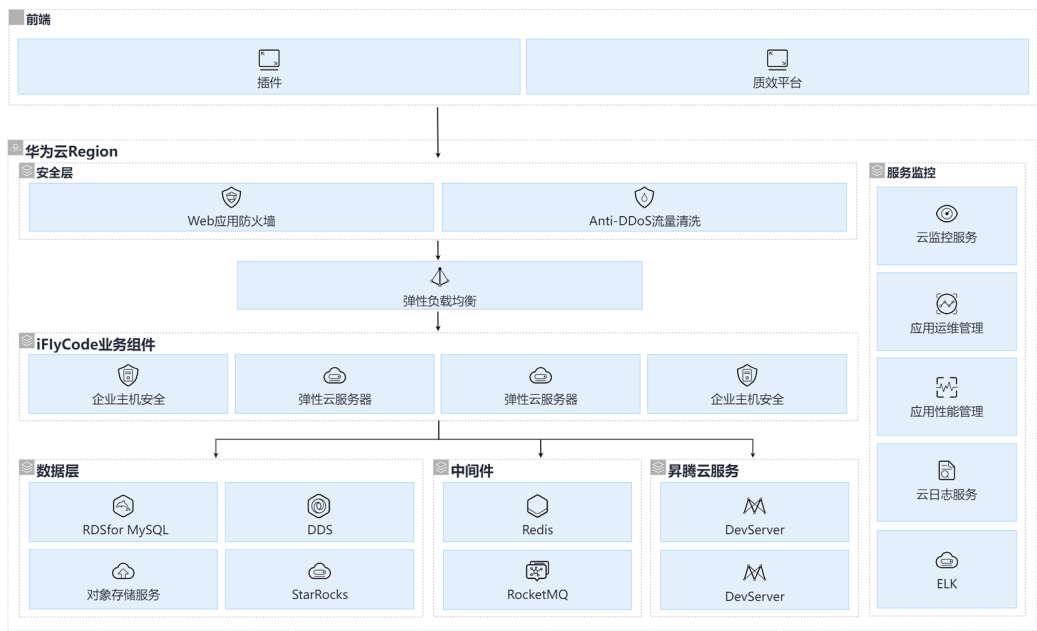


方案主要由华为云昇腾云服务+星火飞码（iFlyCode），基于讯飞星火代码大模型、推理服务、Embedding模型，提供全流程的智能编程助手能力。

- 以IDE插件的产品方式，提供各项产品能力，包括智能问答、助理智能体、辅助编程、辅助开发等。服务于互联网软件行业、能源、工业汽车、金融等各个行业；
- 除了插件的使用方式，也提供OpenAPI的方式，供用户自由灵活的接入使用；
- 效能管理平台，协助客户分析使用情况，包括用户分析、效能度量指标的分析，让使用的结果可度量，可量化。

部署架构图

图 1-2 部署架构图



方案通过华为云昇腾云/DDS/RDS/RocketMQ等实现智能编程助手能力：

- 通过昇腾云服务的算力进行星火大模型推理。
- 通过RDS存储结构化的业务数据。
- 通过DDS存储用户问答数据、使用数据等，然后同步到StarRocks进行报表呈现
- 通过Redis缓存用户登录状态等信息。
- 通过RocketMQ存储业务消息队列。
- 通过StarRocks存储用户使用数据分析结果，用于报表呈现。

方案优势

- **核心效果优异**：基于优异的讯飞星火大模型和华为昇腾云服务，经过开源的测试集测评，核心场景能力优先，代码生成采纳率可达52%；
- **独特场景功能**：除基础代码辅助功能外，还能提供批量单元测试能力，支持动态修改用例并自动生成测试代码，行覆盖率达80%，人机协同更智能。结合企业代码仓库，智能评审系统可显著提升代码质量与规范性；
- **原生适配自研底座**：全国首个自研算力集群训练，在华为AI算力平台的训练效率可达英伟达A100的90%，性能优先；

- **面向客户价值：**对于最终客户，通过集成RDS/Redis/DDS/RocketMQ等云服务，替代虚拟机开源组件自建形式，方案整体成本节省3500+元/年。

2 资源和成本规划

表 2-1 资源和成本规划

云资源	规格	数量	每月费用 (元)
昇腾云	910B4服务器280T（8卡）	2	2112552
ECS	X86计算 通用计算增强型 c7.4xlarge.2 16核 32GB; CentOS CentOS 7.9 64bit; 通用型SSD 40GB; 通用型SSD 300GB;	2	37640
ELB+EIP	共享型负载均衡(性能保障模式); 1个 全动态BGP; 全动态BGP 带宽 10Mbit/s;	1	6650
Flexus云数据库RDS	MySQL 8.0 主备 轻量型 4核8GB; SSD云盘 240GB;	1	8000
Redis	基础版 5.0 主备 X86 DRAM 副本数2 4GB;	1	2776
DDS	副本集 三节点 通用型 8核16GB; 超高IO 200GB;	1	25500
RocketMQ	rocketmq.b2.large.4; 通用型SSD 200GB;	1	15050
OBS	对象存储 标准存储多AZ存储包 1TB;	1	1161
LTS	X86计算 通用计算增强型 c7.2xlarge.2 8核 16GB; CentOS CentOS 7.9 64bit; 通用型SSD 40GB; 通用型SSD 200GB;	1	2518.22
ECS	X86计算 通用计算增强型 c7.4xlarge.4 8核 32GB; CentOS CentOS 7.9 64bit; 通用型SSD 40GB; 通用型SSD 200GB;	1	12176
WAF	入门版;	1	990
Anti-DDoS	免费	1	0

3 实施步骤

3.1 IFlyCode-VPC管理后台（面向管理员）

3.2 iFlyCode插件（以IDEA为例）

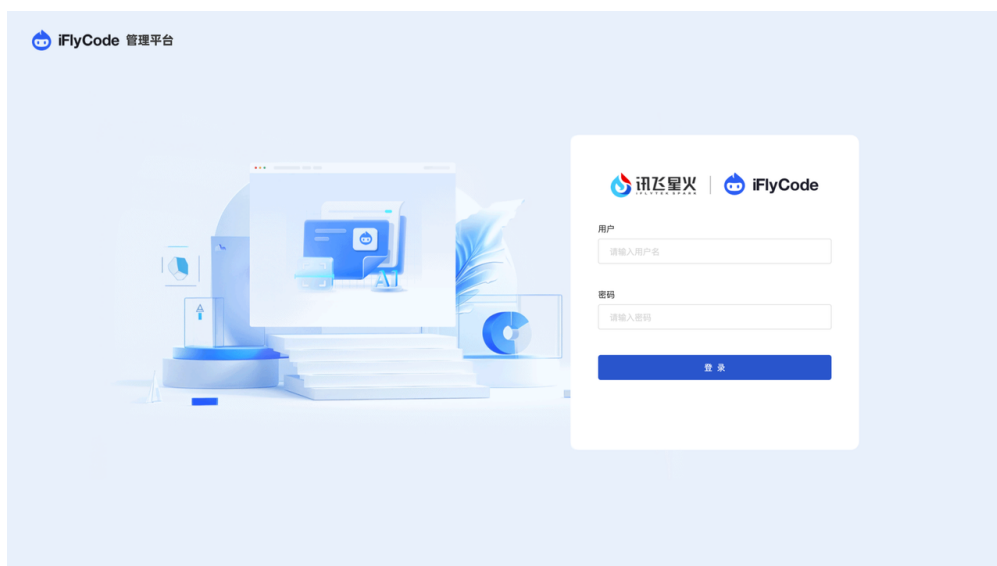
3.1 IFlyCode-VPC 管理后台（面向管理员）

IFlyCode-VPC管理后台面向管理员提供登录激活、插件下载、数据看板、组织管理、成员管理、意见反馈、企业信息等功能，面向普通用户仅提供插件下载功能。

登录激活

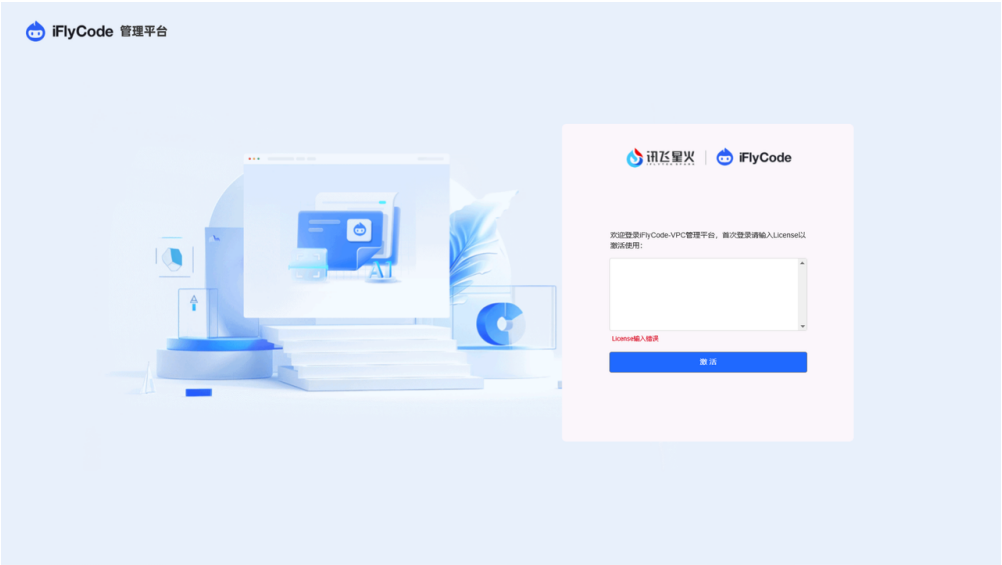
步骤1 登录iFlyCode-VPC管理后台。

图 3-1 登录



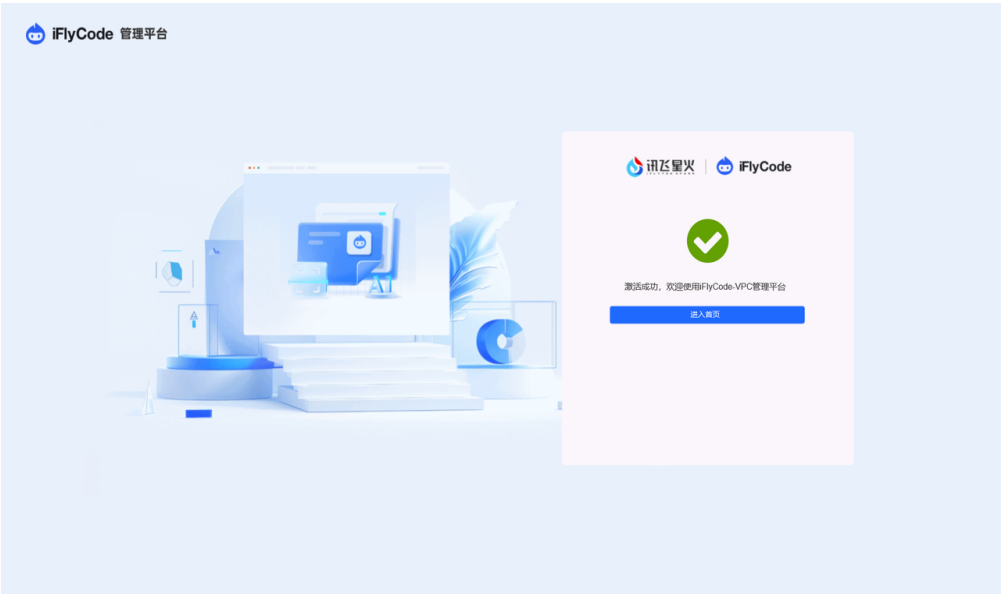
步骤2 输入license激活。

图 3-2 激活



步骤3 激活成功后进入后台使用。

图 3-3 进入后台

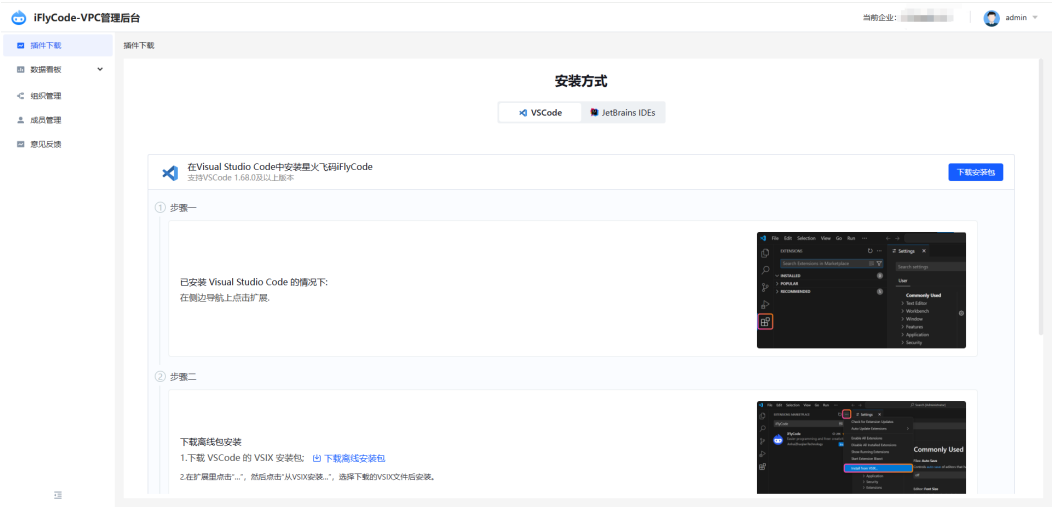


----结束

插件下载

步骤1 单击【插件下载】菜单进入插件下载页面。

图 3-4 插件下载



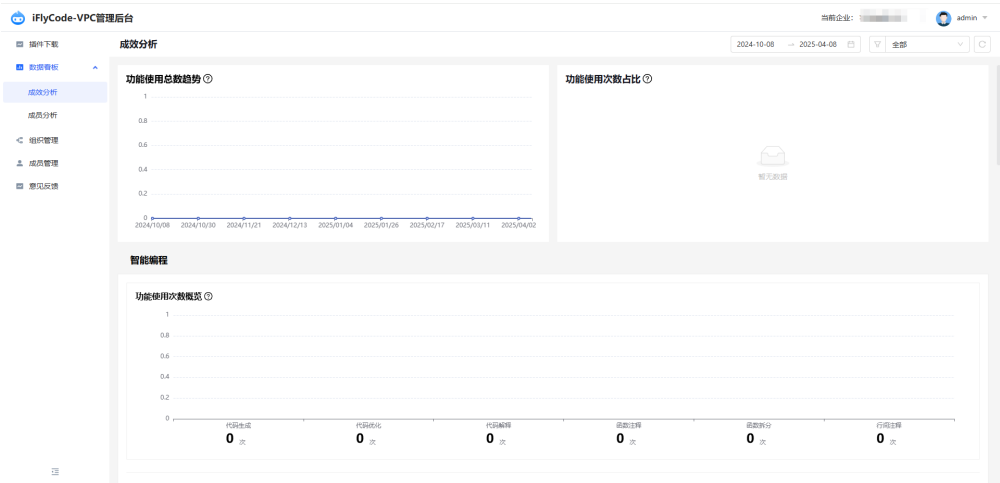
步骤2 选择IDE版本，下载对应插件包。

----结束

查看数据看板-成效分析

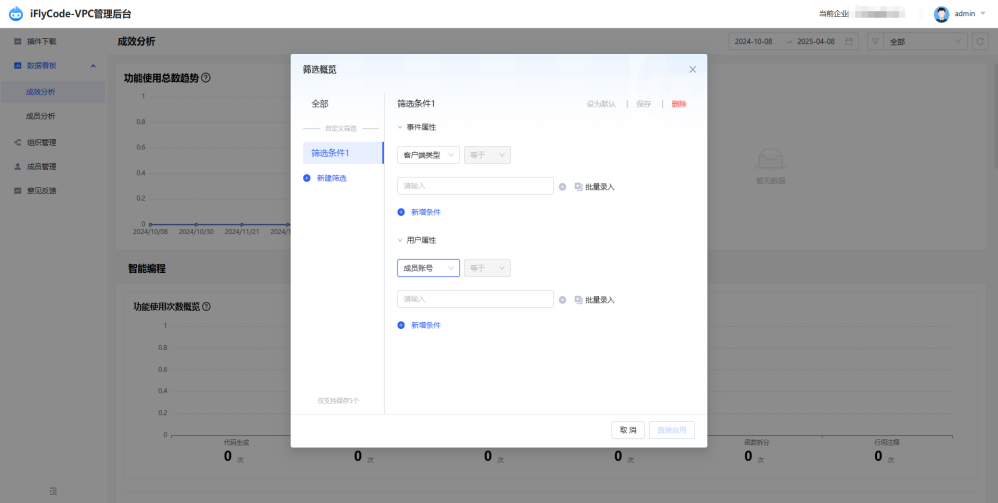
步骤1 单击【数据看板-成效分析】菜单进入数据看板-成效分析页面，查看对应维度数据。

图 3-5 数据看板-成效分析



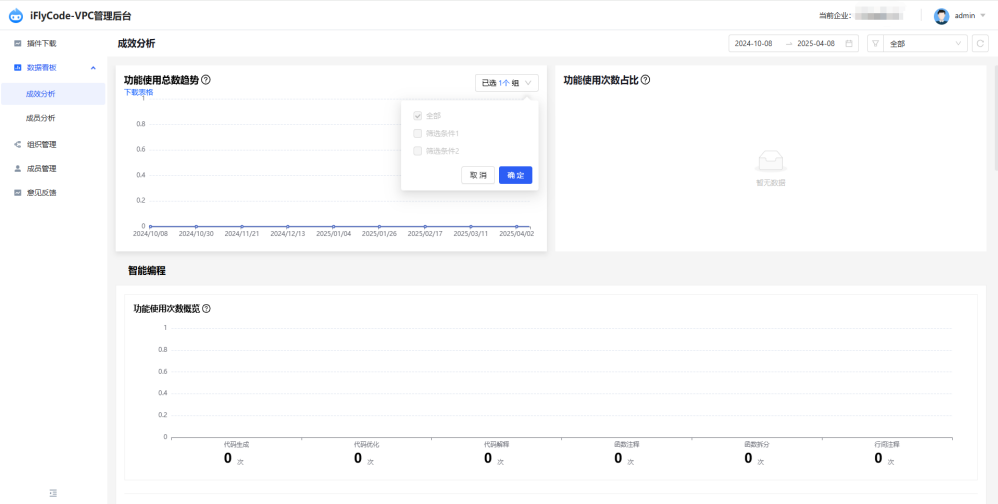
步骤2 单击右上角筛选器按钮可以根据条件创建筛选器。

图 3-6 创建筛选器 1



步骤3 单击具体指标可以添加对比，最多支持3个筛选器。

图 3-7 创建筛选器 2



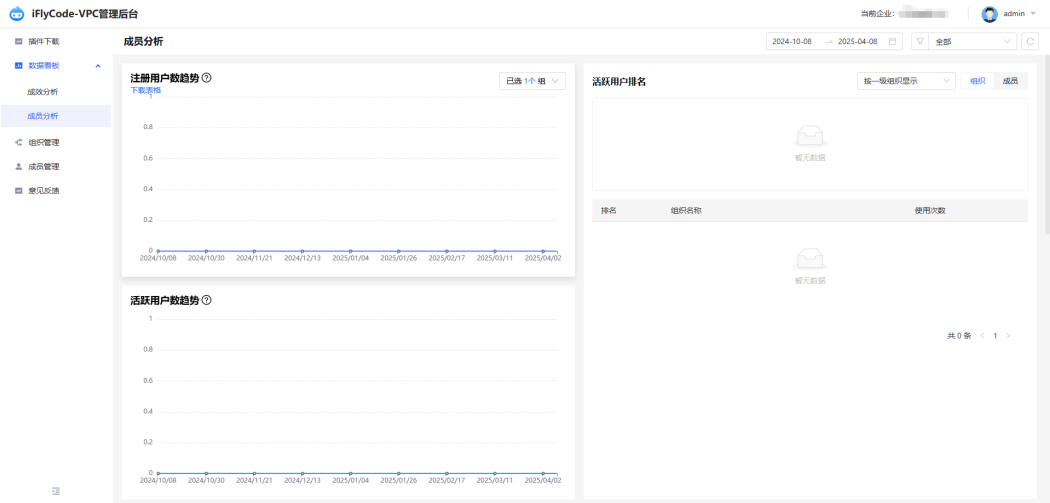
步骤4 单击具体指标可以下载表格进行数据导出。

----结束

查看数据看板-成员分析

单击【数据看板-成员分析】菜单进入数据看板-成员分析页面，查看对应维度数据。

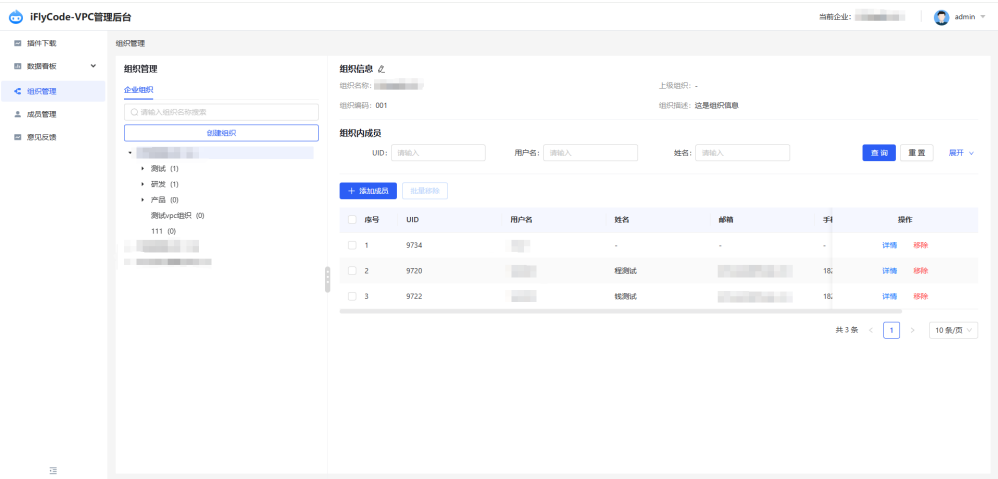
图 3-8 数据看板-成员分析



组织管理

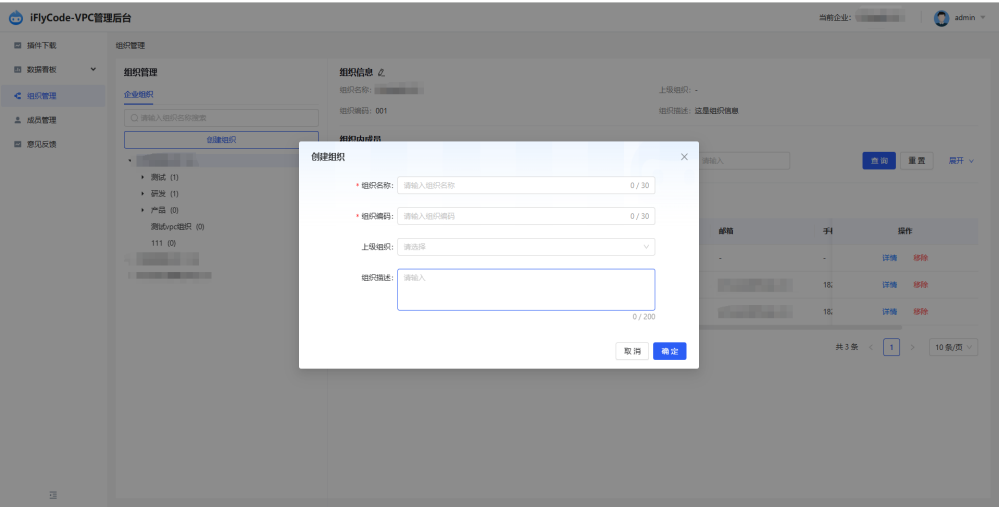
步骤1 单击【组织管理】菜单进入组织管理页面，查看组织和组织下成员信息。

图 3-9 组织管理



步骤2 单击创建组织，填写组织名称等信息，填写完单击确定生成组织。

图 3-10 创建组织



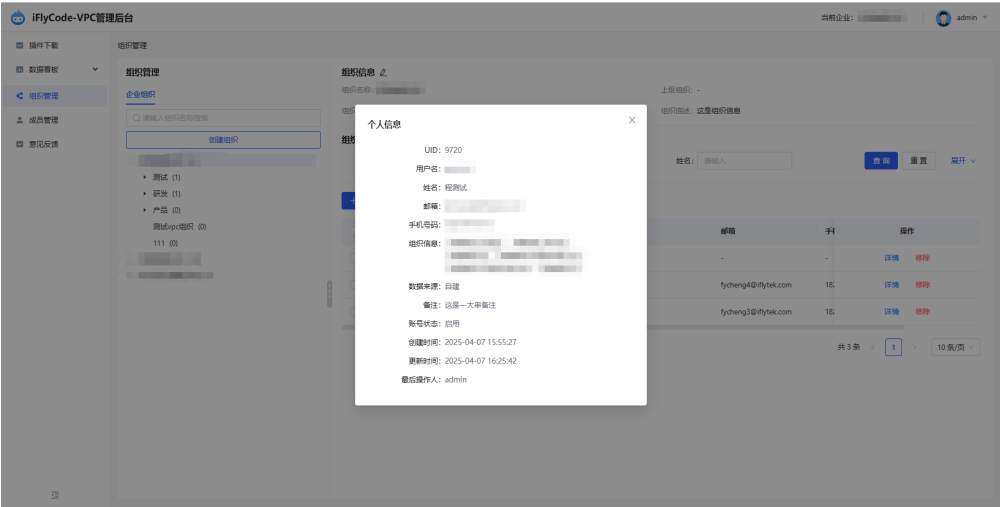
步骤3 单击添加成员，可以从其他组织节点中选择成员挂载到该节点上。

图 3-11 添加成员



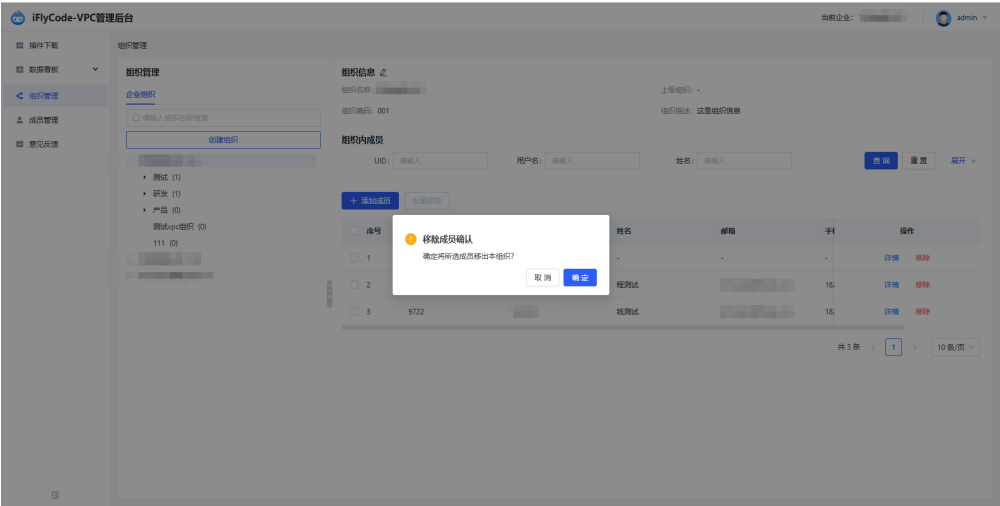
步骤4 单击详情可以查看成员详细信息。

图 3-12 查看成员详细信息



步骤5 单击移除可以将成员从该节点中移除。

图 3-13 节点移除

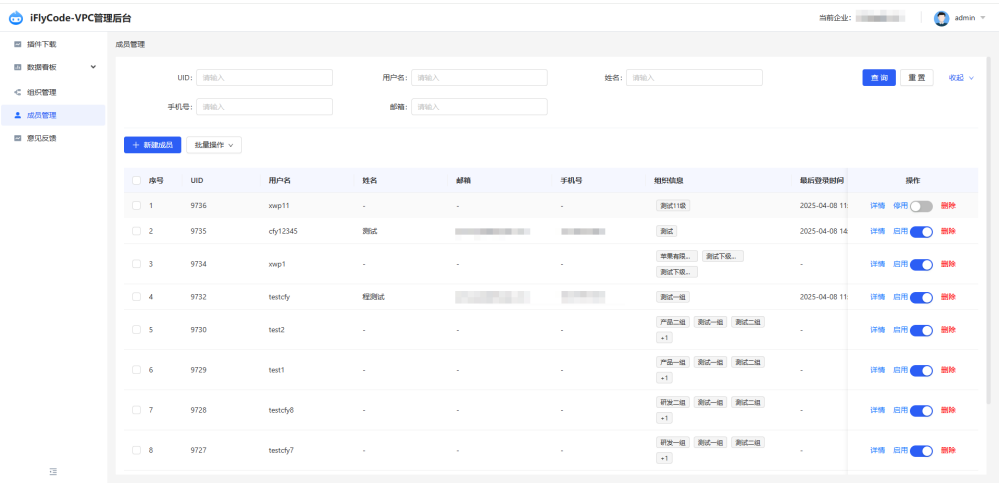


----结束

成员管理

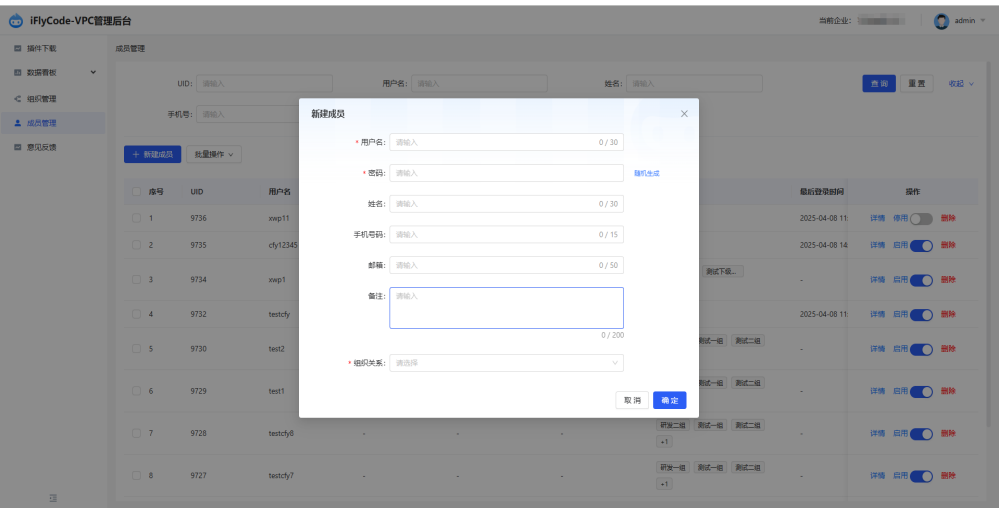
步骤1 单击【成员管理】菜单进入成员管理页面，支持新增、编辑、查看、删除成员信息。

图 3-14 成员管理



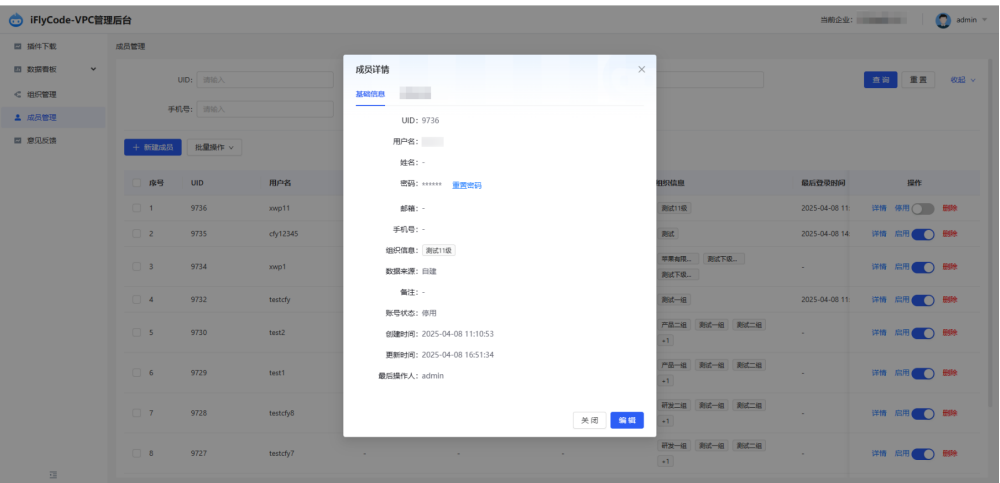
步骤2 单击新建成员，填写完用户名等信息后，单击确定按钮完成成员账号生成。

图 3-15 新建成员



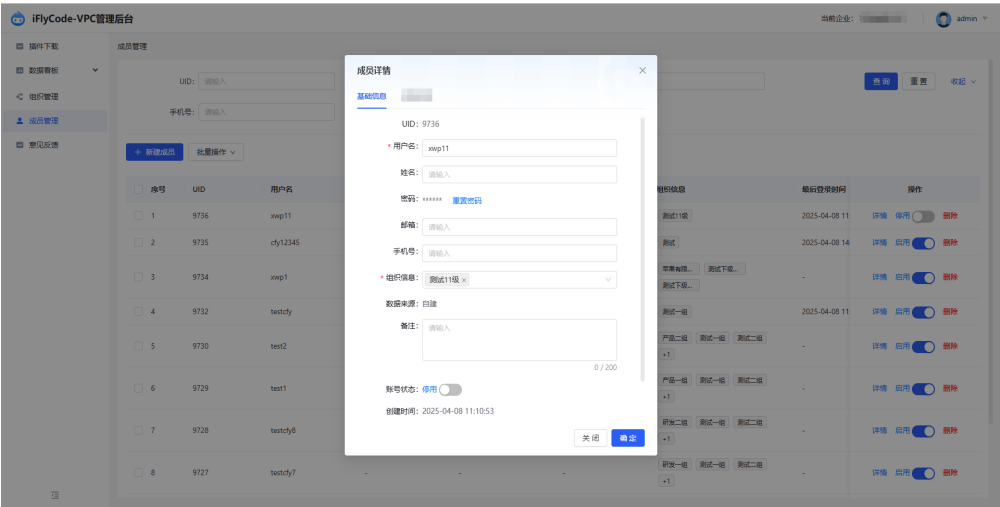
步骤3 单击详情可以查看成员信息。

图 3-16 查看成员信息



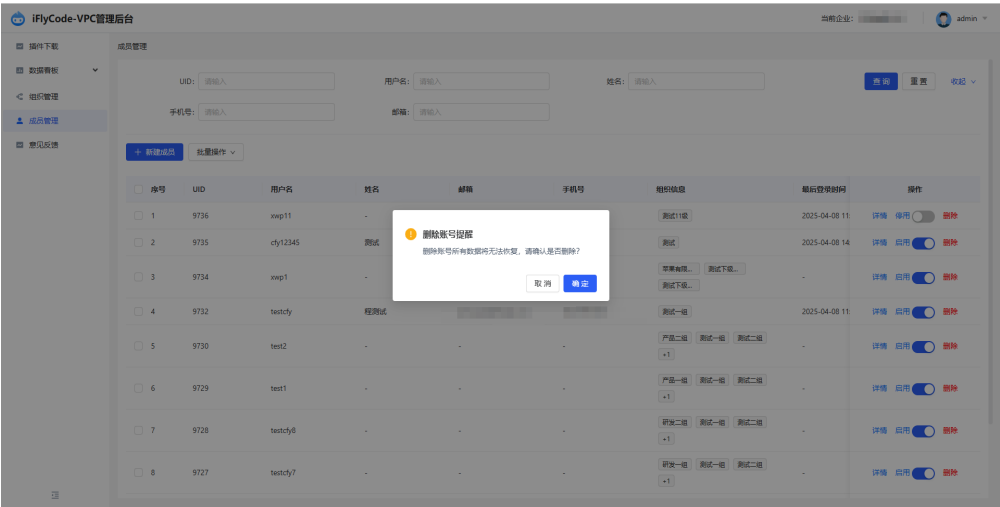
步骤4 在详情中单击编辑，可以修改成员信息。

图 3-17 修改成员信息



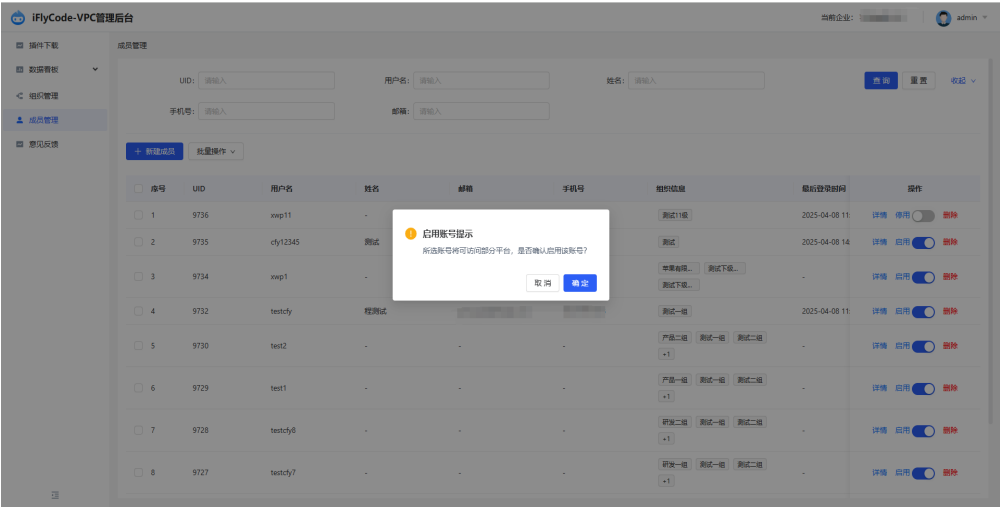
步骤5 单击删除可以删除成员账号。

图 3-18 删除成员账号



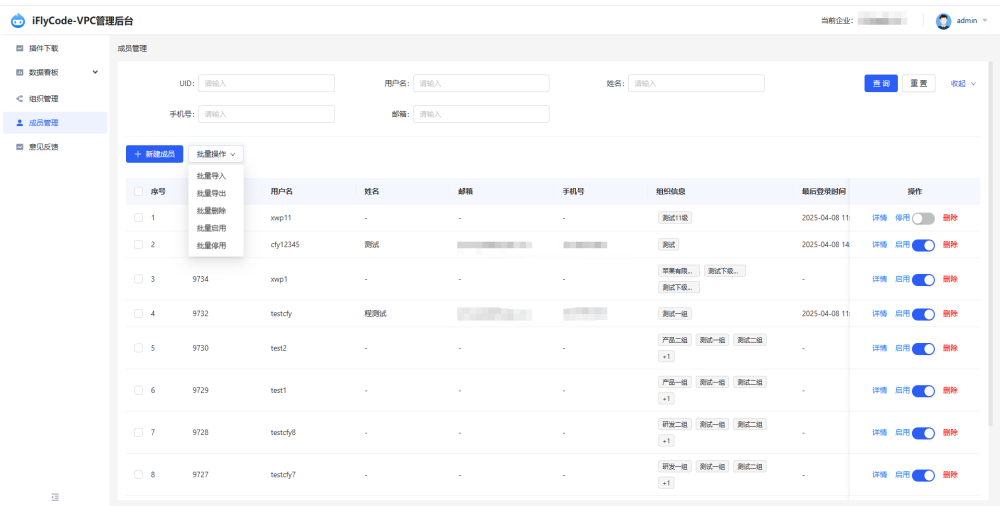
步骤6 单击启/停用开关可以启用或停用成员账号。

图 3-19 启用或停用成员账号



步骤7 单击批量操作按钮，可以批量导入、批量导出、批量删除、批量停用、批量启用成员账号。

图 3-20 批量操作



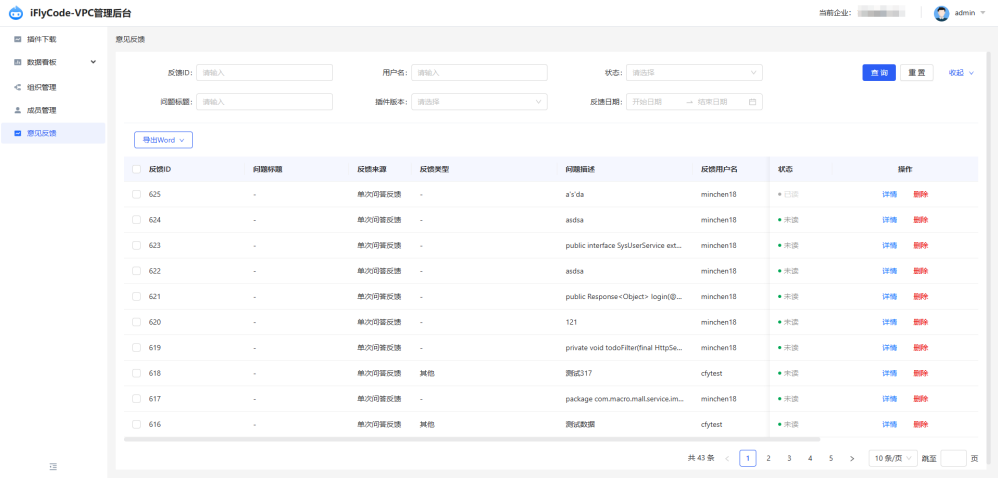
步骤8 支持按用户名、姓名、手机号、邮箱、UID等搜索成员账号。

----结束

意见反馈

步骤1 单击【意见反馈】菜单进入意见反馈页面，查看成员使用过程中反馈的信息。

图 3-21 意见反馈



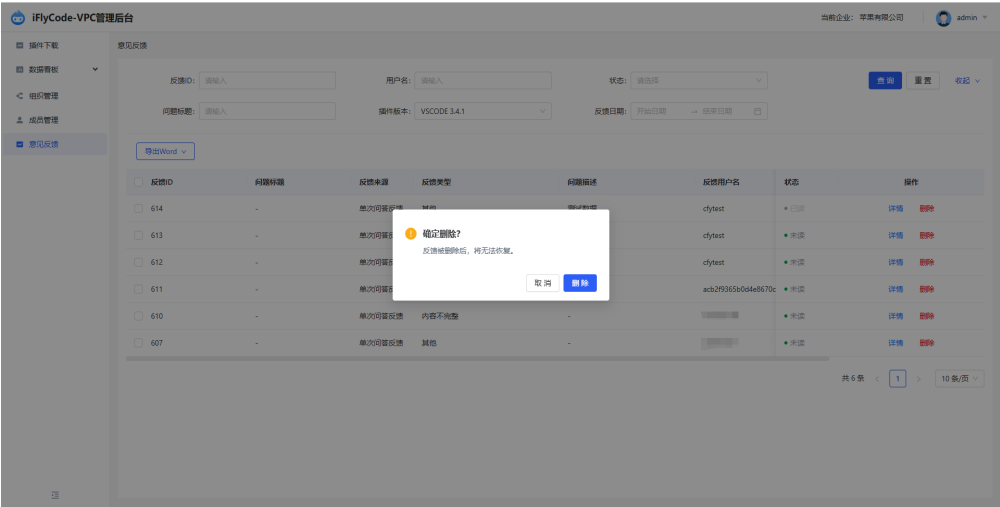
步骤2 单击详情可以查看意见反馈的详细信息。

图 3-22 反馈详情



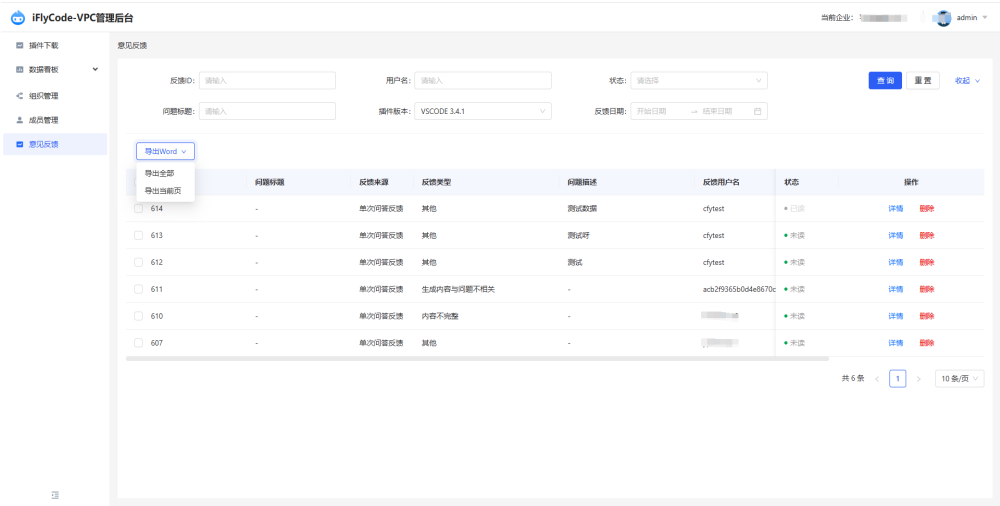
步骤3 单击删除可以删除反馈信息。

图 3-23 删除反馈信息



步骤4 单击左上角导出按钮可以导出当前页反馈或全部反馈信息。

图 3-24 导出反馈信息



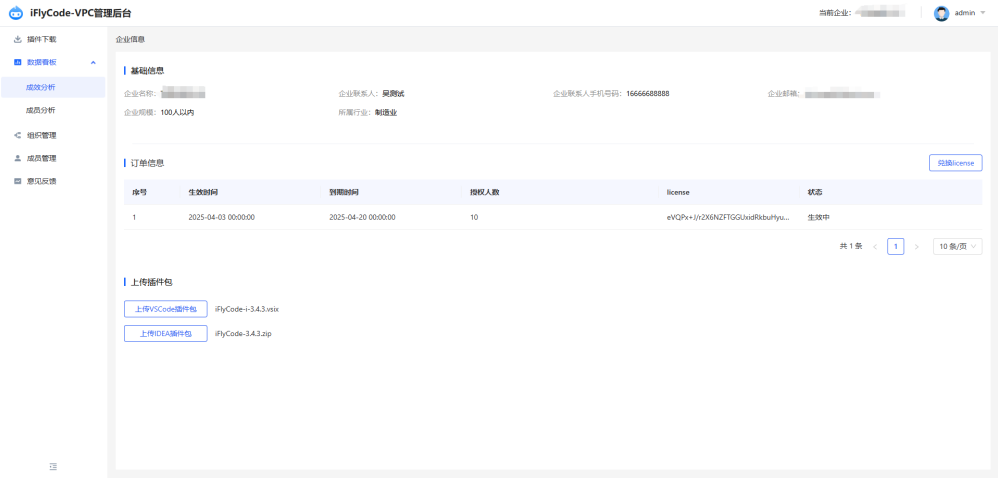
步骤5 支持按反馈ID、用户名、状态、标题、插件版本、反馈日期条件查询反馈信息。

----结束

企业信息

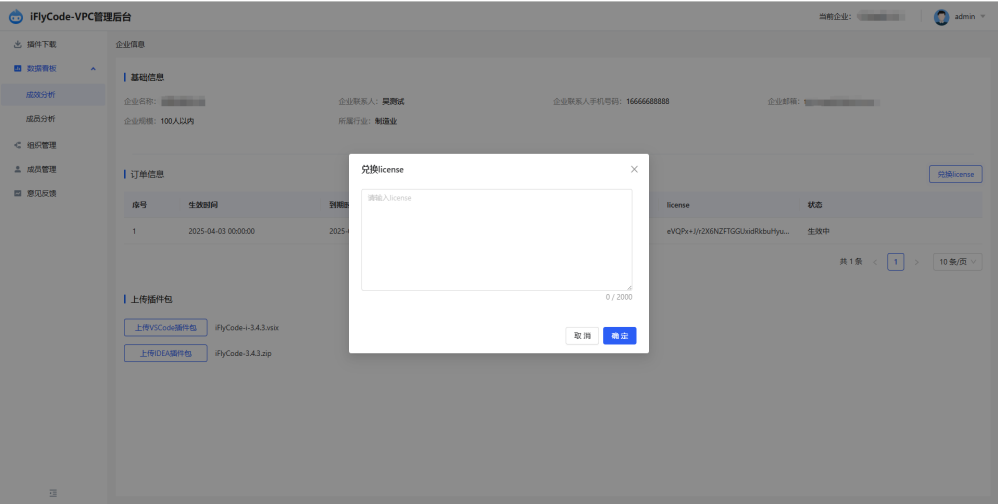
步骤1 鼠标移到右上角头像图标，单击【企业信息】菜单，可以查看企业信息。

图 3-25 企业信息



步骤2 单击兑换license，可以用于权益临期时再次激活产品。

图 3-26 兑换 license



步骤3 单击上传插件包，可以上传不同IDE版本的插件包用于更新。

----结束

3.2 iFlyCode 插件（以 IDEA 为例）

下载和安装插件

- 步骤1 用户登录iFlyCode-VPC管理后台，在插件下载页面下载插件包。
- 步骤2 单击设置-插件-从磁盘安装插件，选择插件包安装，安装完成后重启IDE生效。

图 3-27 下载和安装插件

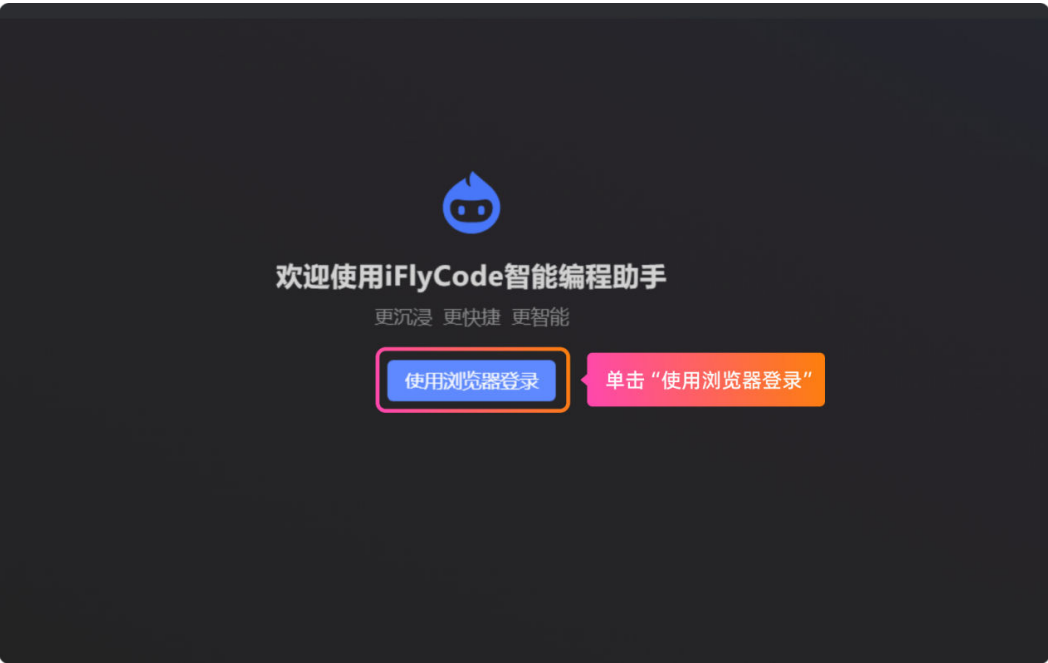


----结束

登录插件

单击使用浏览器登录，跳转登录页面，输入账号密码完成登录

图 3-28 登录插件



核心功能介绍

代码生成：iFlyCode支持在编辑器内根据注释、函数名自动生成代码，同时也支持在问答窗口通过自然语言描述功能需求，自动生成代码，且支持单行模式和智能模式；

图 3-29 代码生成



代码补全：iFlyCode支持在编辑器内，通过方法名、上下文等信息触发代码补全，支持单行和多行补全，单行模式下仅生成一行代码建议，推荐效果和响应速度更好，智能模式下通过实时分析代码的上下文，智能生成单行或多行补全代码建议。同时支持跨文件感知能力。在已有上下文的代码中，回车、空格均会自动触发代码建议，使用Tab键快捷采纳建议，Esc拒绝建议。

图 3-30 代码补全



代码解释：iFlyCode支持对选中代码进行代码解释，包括该段代码的作用和含义、代码中的类和方法等。选中代码后，右键选择“代码解释”使用代码解释功能。同时支

持在代码段上方单击“代码解释”快捷选项，代码解释同时也支持对反馈的结果进行点踩/点赞/意见反馈。

图 3-31 代码解释 1

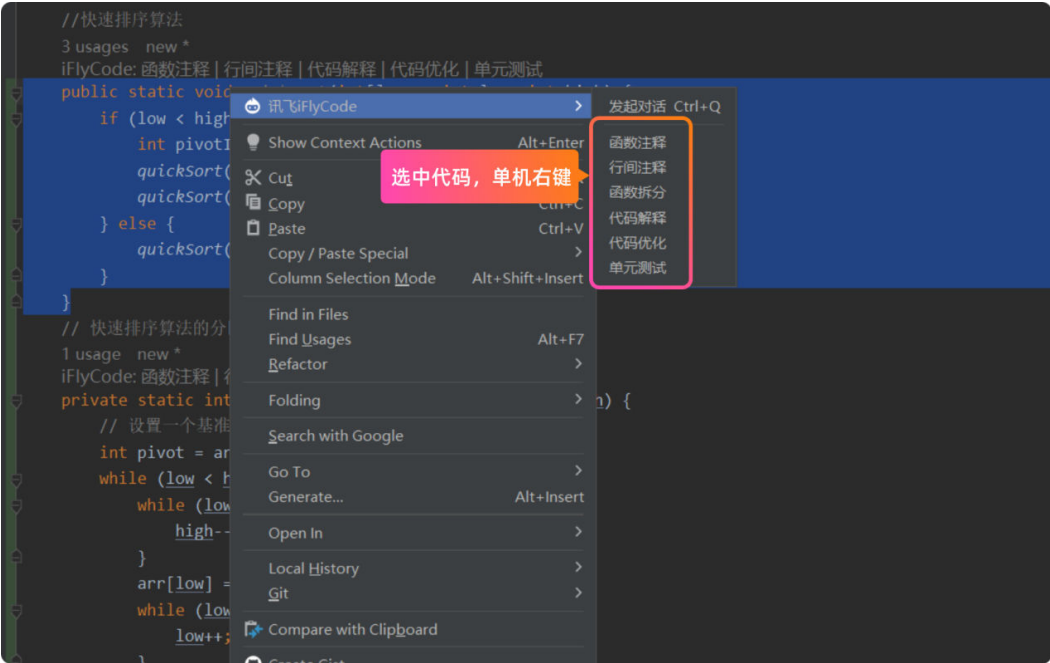
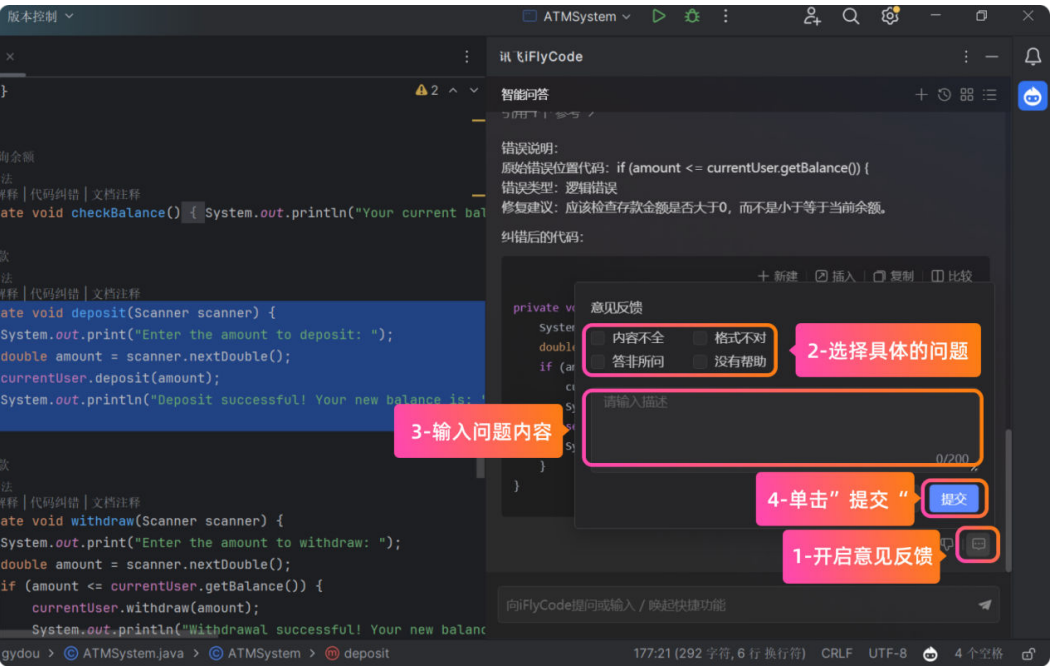


图 3-32 代码解释 2



图 3-33 代码解释 3



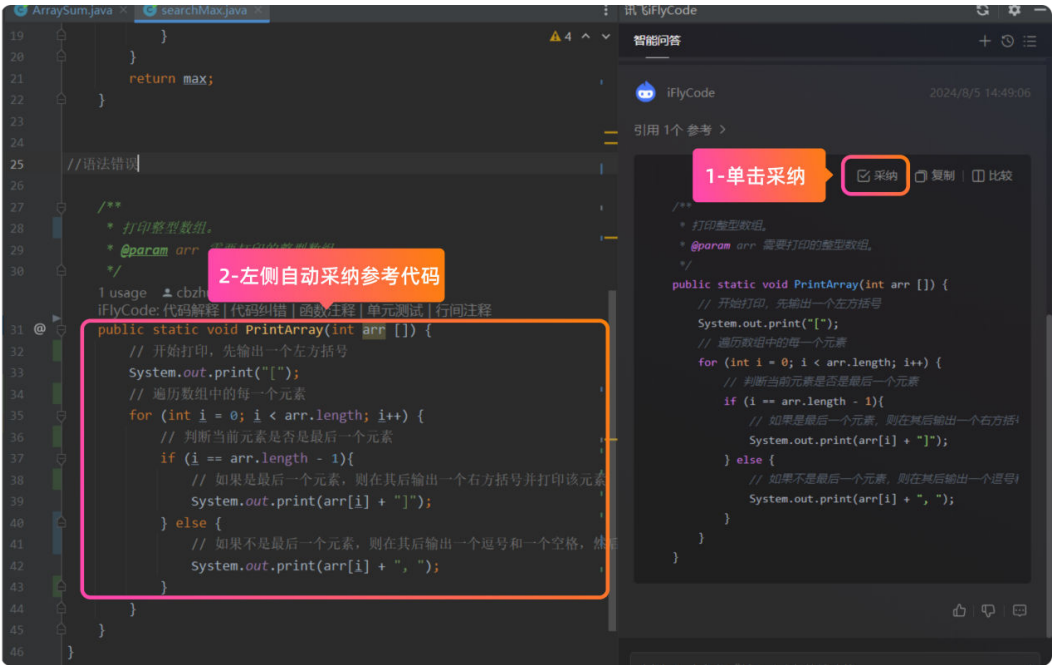
函数注释：iFlyCode支持对选中的函数生成函数注释。选中代码后，右键选择“函数注释”使用函数注释功能；并且支持在代码段上方单击“函数注释”快捷选项，同时支持对反馈的结果进行点踩/点赞/意见反馈。另外函数注释支持对反馈的注释进行一键采纳，方便便捷，同时支持复制功能，插入到任意位置。函数注释支持批量，批量函数注释可以对文件中的方法进行注释，并且支持一键采纳，选中文件单击鼠标右键即可触发批量函数注释

图 3-34 函数注释



行间注释：iFlyCode支持对选中的函数生成行间注释。并且支持在代码段上方单击“行间注释”快捷选项，同时支持对反馈的结果进行点踩/点赞/意见反馈。另外行间注释支持对生成的注释进行一键采纳，方便便捷，同时支持比较功能，选中该行代码单击比较，进行逐行采纳。

图 3-35 行间注释



函数拆分：iFlyCode支持对选中代码进行函数拆分，当函数大于等于20行，且少于12000个字符时，可以将函数分解为更小、更短的易于理解的段函数。选中代码后，右键选择“函数拆分”使用函数拆分。

SQL生成/优化：iFlyCode支持配置MySQL、PostgreSQL、OceanBase、达梦数据源，结合数据库结构进行SQL生成和SQL优化。以MYSQL为例：首先进入工具箱，单击SQL生成/优化板块，输入必要信息，链接测试并保存。

图 3-36 SQL 生成/优化 1

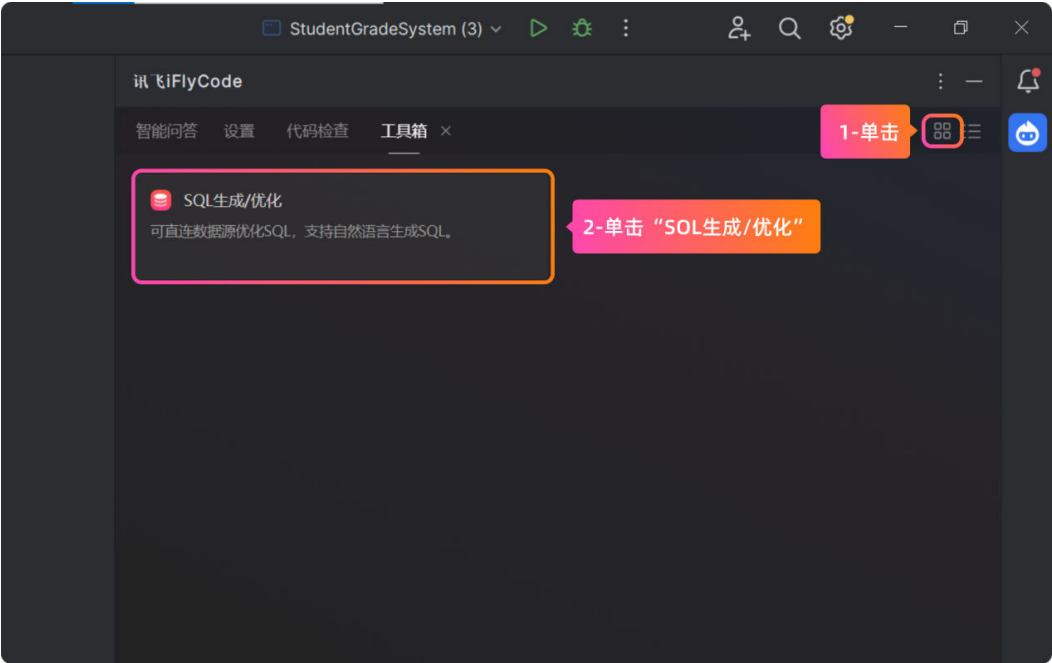
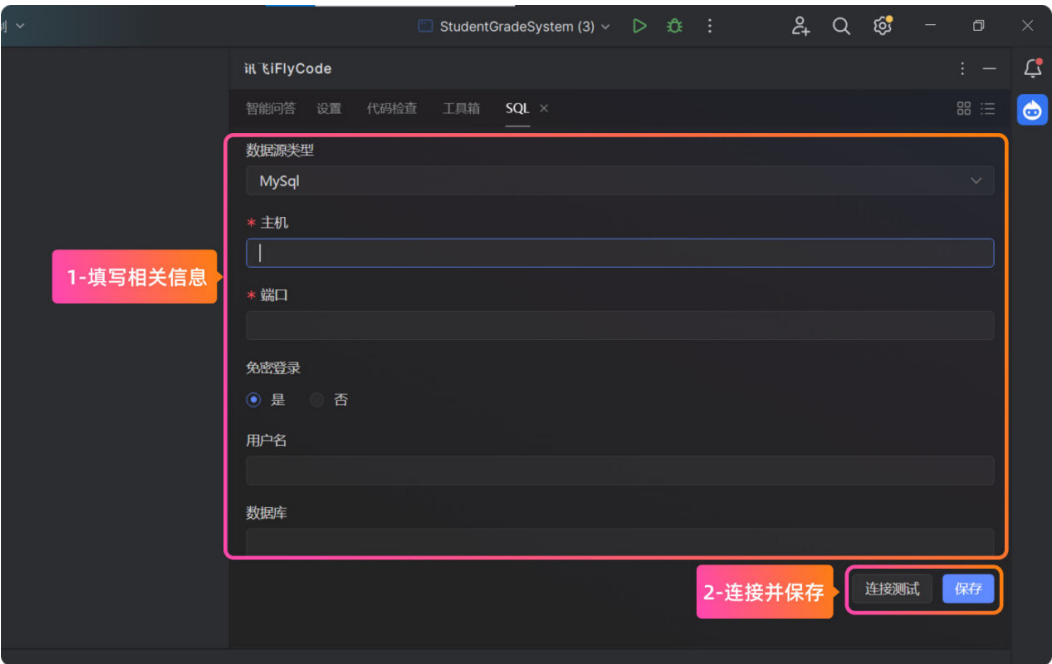


图 3-37 SQL 生成/优化 2



配置完成后，选择数据库和数据表，输入框通过“/”唤起SQL生成或者SQL优化功能，对SQL进行生成和优化。

图 3-38 SQL 生成/优化 3

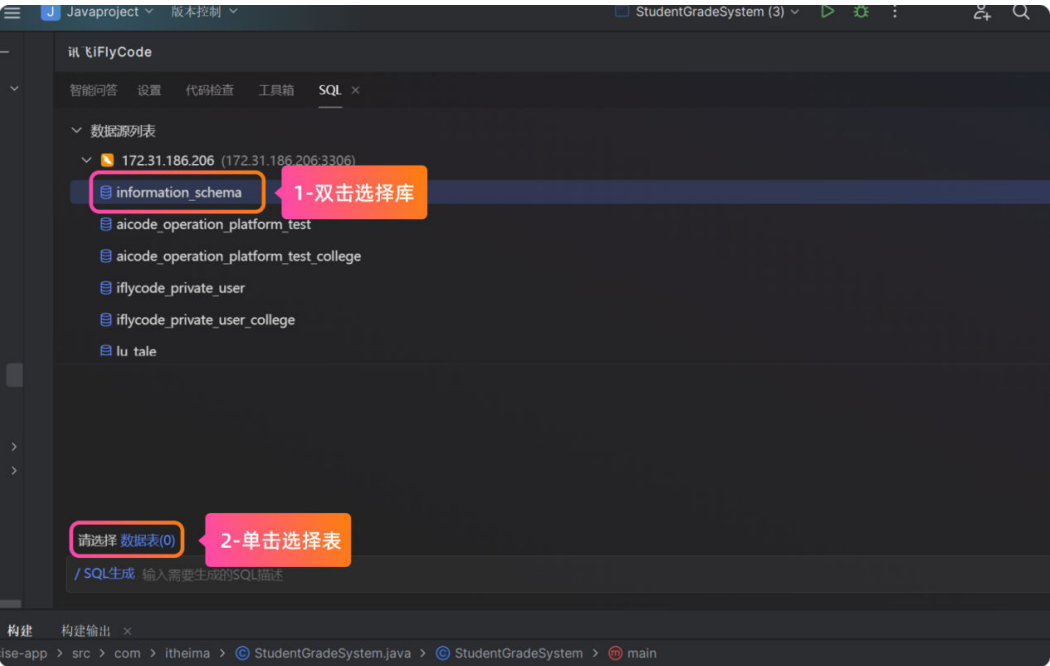


图 3-39 SQL 生成/优化 4

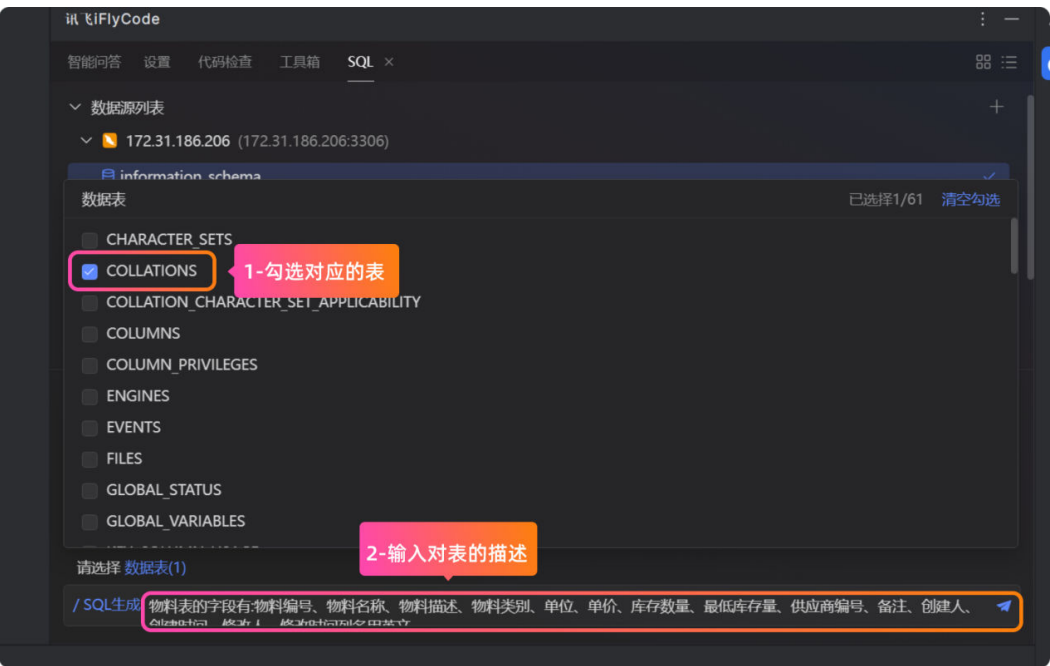
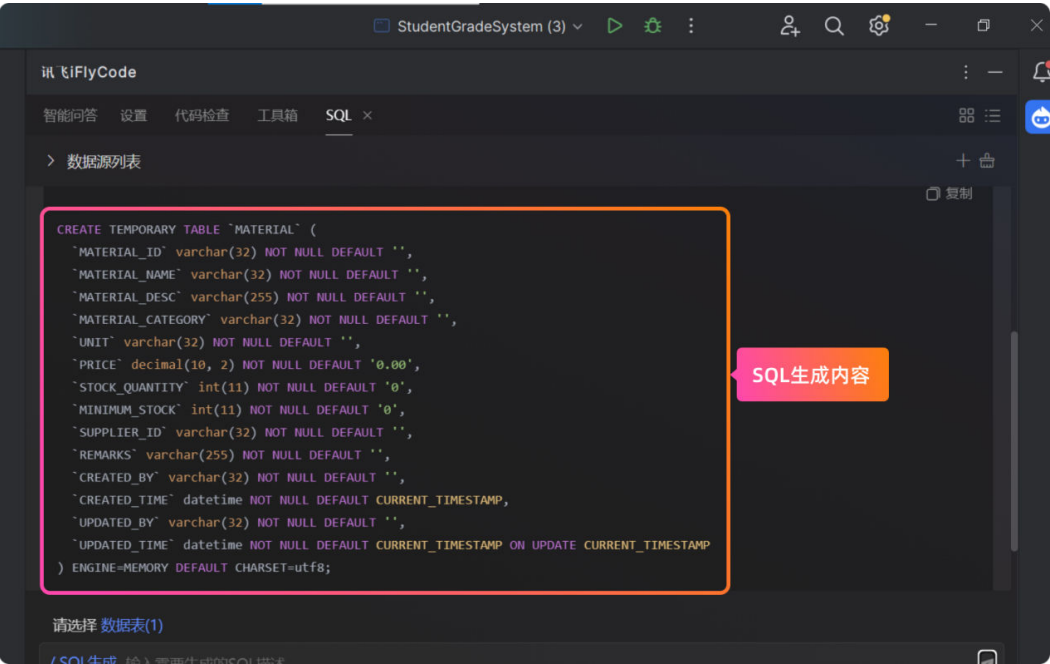
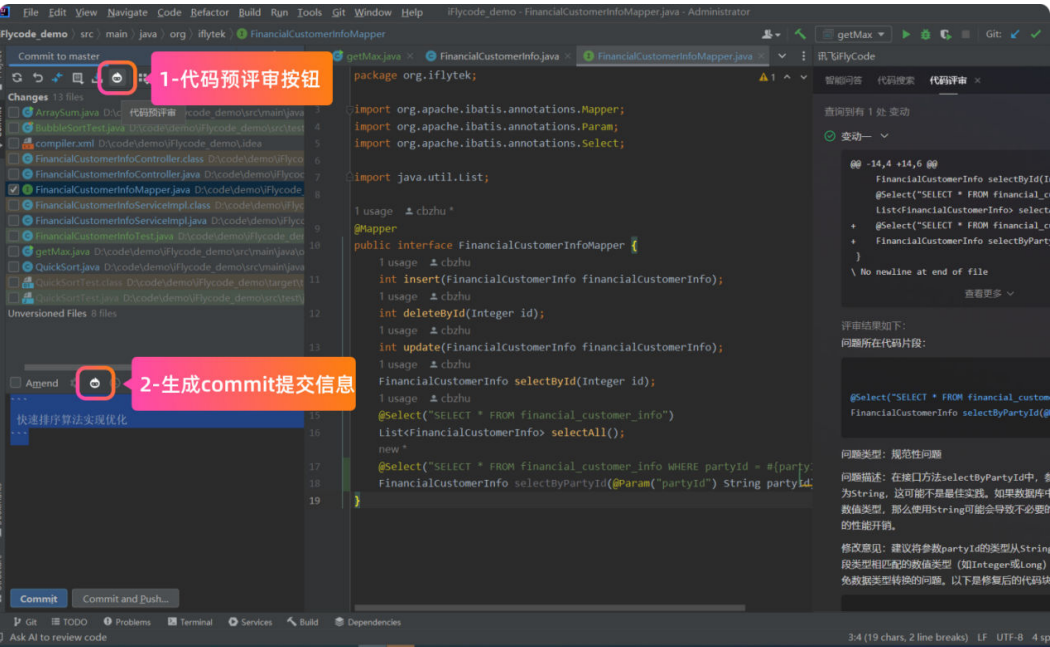


图 3-40 SQL 生成/优化 5



代码预评审：提交工作空间内的源代码后，在IDEA侧边栏面板中，单击iFlyCode图标按钮，iFlyCode通过理解您暂存的更改文件，对提交的代码进行代码预评审，检查代码的完整性、逻辑性和安全性等问题，且支持一键生成Commit提交信息。

图 3-41 代码预评审



研发自由问答：iFlyCode提供基于上下文的多轮对话智能问答功能，您可以直接在 IDE 对话框中输入研发过程中遇到的相关问题，iFlyCode将快速准确地返回答案，无需跳转至其他页面，旨在提供更好的编程体验。

图 3-42 研发自由问答



对话式编程：在编码过程中，您可以在对话框中描述编码任务和要求，iFlyCode会结合项目信息和您的输入生成可用的代码。

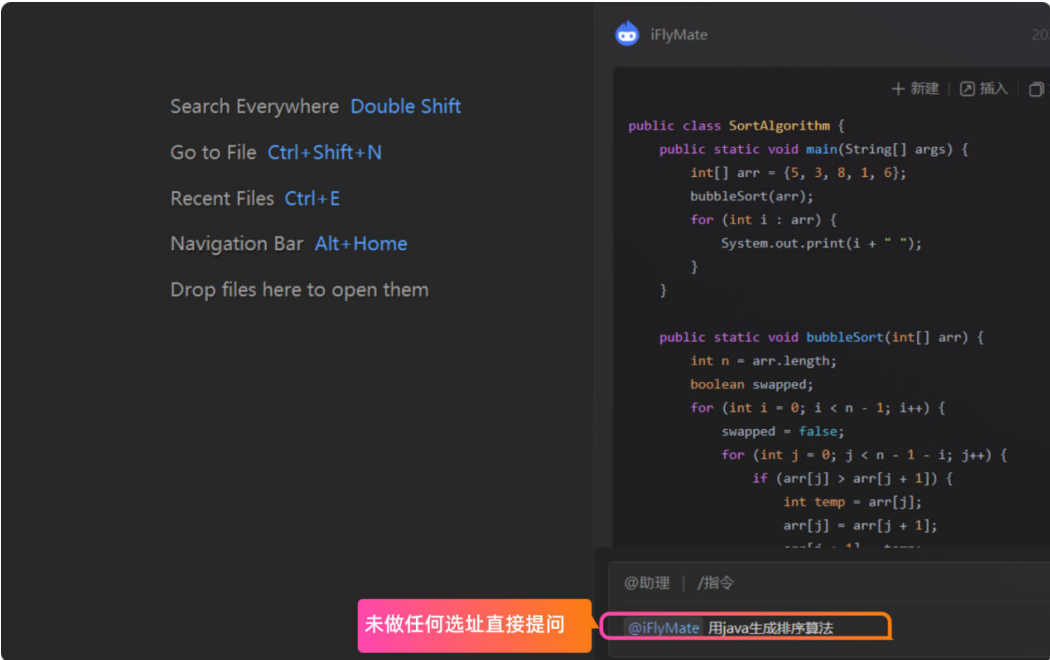
图 3-43 对话式编程



助理问答：iFlyCode支持唤起具有领域知识和专业技能的软件开发助理（如iFlyMate、iFlyTest），通过各项助理能力，协助更多软件开发及相关人员，快速有效地解决更多痛点问题。

智能问答时，iFlyMate可以不做任何选择直接进行提问，插件会自动加载该配置；另外iFlyMate支持指令模式，指令包含函数注释、函数拆分、代码解释、代码优化，在使用时选择指令并选中代码即可。

图 3-44 助理问答



用户反馈：iFlyCode支持提交意见反馈，用户可通过单击意见反馈按钮，提交自己的使用反馈。

图 3-45 用户反馈



4 修订记录

表 4-1 修订记录

发布日期	修订记录
2025-05-09	第一次正式发布。